University of Bath

PHD

From Recognition to Susceptibility: Functional characterization of Plant- specific LIM-domain containing proteins in plant-microbe interactions

Parkes, Toby

Award date:
2020

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy
If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Download date: 04. May. 2023

From Recognition to Susceptibility: Functional characterization of Plant-specific LIM-domain containing proteins in plant-microbe interactions

Volume 1/3

by

Toby Parkes

Thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

March 2020

Declarations

Copyright notice

Attention is drawn to the fact that copyright of this thesis/portfolio rests with the author and copyright of any previously published materials included may rest with third parties. A copy of this thesis/portfolio has been supplied on condition that anyone who consults it understands that they must not copy it or use material from it except as licenced, permitted by law or with the consent of the author or other copyright owners, as applicable.

Declaration of authorship

I am the author of this thesis, and the work described therein was carried out by myself personally, with the exception of chapter 3 where 30% of the experimental work was carried out by Professor Eric Holub and Dr Volkan Cevik and chapter 6 where formulation of some of the ideas and 5% of the experimental work was performed in collaboration with Prof Ksenia Krasileva and Erin Baggs.

Table of Contents

Declarations **3**
Copyright notice 3
Declaration of authorship 3
List of Figures **8**
List of Tables **12**
Acknowledgements **13**
Abstract **14**
Abbreviations **16**
Chapter 1 ..... **19**
General Introduction **19**
Global perspective on land usage: Food security, anthropomorphic population growth and climate change 19
Current knowledge of the plant immune system 21
Responses to invasion and the classic model of plant immunity 21
Cell surface receptors and the first layer of plant innate immunity 23
Suppression of MTI by invading pathogens 24
The second layer of plant immunity: detecting the effector 26
NLRs throughout the Kingdoms 29
On the Origin of NLRs 31
NLR-like proteins and the diversity of NLR architectures 32
Immune signalling events in phyto-NLR systems 33
Structural insights into NLR activation 39
Pairing up – the emergence of dual NLR systems and the sensor-helper model 41
Integrated domains: How to snare an effector 42
The LIM domain as an integrated domain 43
Albugo candida and white rust resistance 44
Thesis Aims 46
Chapter 2 ..... **47**
Materials and Methods **47**
Biological Material 47
Trypan Blue staining 48
Cloning of *WRR5A* and *WRR5B* 48
USER Cloning 49
Floral dipping 50
DNA extraction and Genotyping 51

RT-qPCR 52
Transient expression 52
Protein sub-cellular localisation and microscopy 53
Protein extraction 53
Microsomal and Nuclear Fractionations 54
Immunoprecipitation 55
Immunoblotting 56
IP-MS 57
EMS mutagenesis and Mutant analysis 57
Bioinformatics 58
EMS mutagenesis candidate gene identification 58
Gene ontology analysis 58
RNA sequence analysis 58
Phylogenetic analysis – sequence alignments and tree building 60
Synteny analyses 60
Chapter 3 69
Identification of multiple resistance genes conferring immunity to *Albugo candida* isolate AcEM2 in the *Arabidopsis thaliana* accession Col-0 69
Introduction 69
Results 73
Genetic analysis reveals multiple independent *White Rust Resistance* genes in the *Arabidopsis thaliana* Columbia accession 73
Characterization of distinct resistance phenotypes mediated by different *White Rust Resistance* genes in *Arabidopsis thaliana* accession Columbia 77
Cloning of resistance genes in the *WRR5* locus 79
An intact P-loop in both WRR5A and WRR5B is required to cause an immune response 82
DAR5 is the underlying gene for *WRR7* mediated resistance 85
Discussion 89
Chapter 4 93
Mechanistic insights into *WRR5A* and *WRR5B* mediated immunity 93
Introduction 93
Results 96
Overexpression of *WRR5A* and *WRR5B* stimulates an immune response in *A. thaliana* Ws-2 after infection with AcEM2 96
WRR5A and WRR5B form a heterodimer in *A. thaliana* 99
WRR5A and WRR5B localise to the plasma membrane 102
WRR5A and *WRR5B* mediated immunity requires helper NLRs 106

IP-MS identifies WRR5A and WRR5B interacting proteins 110
Discussion 117
Chapter 5 122
Insights into *WRR7* mediated immunity 122
Introduction 122
Results 128
WRR7 is upregulated in *Arabidopsis thaliana* ecotypes following infection by *Albugo candida* isolates. 128
RNA sequence analysis reveals genes regulated by CAMTA transcription factors during colonisation of *A. thaliana* leaves by *A. candida* 130
Identification of genes involved in *WRR7* immune signalling via EMS mutagenesis. 138
Bulk segregant analysis of susceptible BC F_2 plants reveals novel players in *WRR7* mediated immunity 141
Cloning and complementation of candidate genes in the EMS lines 148
CAMTA2 but not *CAMTA1* or *CAMTA3* are required for *WRR7* immunity 153
Expression analysis reveals that several EMS mutants have impaired *WRR7* expression 156
CAMTA2 interacts with CaM2/3 158
ADR1s and NRGs are not required for *WRR7* mediated immune signalling 162
The C442Y mutation in WRR7 induces autoimmunity 163
The EMS 138 S8F mutation abolishes autoimmunity caused by the WRR7-C442Y mutation 166
Discussion 170
Chapter 6 177
The evolutionary history of *WRR5B* and *WRR7* 177
Introduction 177
Results 182
The DA1 family is divided into two distinct clades 182
Overexpression of DA1 family proteins in *Arabidopsis thaliana* results in phenotypic abnormalities 184
DA1 family proteins are unstable following *A. candida* infection 187
The DA1 family Clade II proteins as decoys 190
LIM-Peptidase domains of WRR5B and WRR7 do not show peptidase activity on EOD1 193
Identification of Resistance genes with integrated LIM-Peptidase domains 196
Clade II of the DA1 family evolved at the base of the Brassicaceae lineage 199
The *Camelina sativa* genome harbours a homolog of *WRR5B* but not *WRR7* 202
Tracking the evolution of WRR5B 204
The evolution of *WRR7* 206

Domain swap experiments of WRR7 and DA1 family LIM-peptidase domains 208
Discussion 211
Chapter 7 215
General discussion 215
Identification of novel *White Rust Resistance* genes 215
WRR5A and WRR5B operate by the sensor-helper model of NLR activation 215
A non-canonical CC_R type resistance protein WRR7 provides resistance against *Albugo candida* 218
Calcium signalling and the phytobiome 222
Chromatin remodelling and the epigenetic regulation of plant immunity 225
ABA signalling in plant immunity 227
The LIM-Peptidase protein family and their role in plant immunity 228
White rust resistance genes in *Arabidopsis thaliana* 230
Evolution of the LIM-Peptidase integrated decoy domain 231
Deploying the integrated domain in crop varieties 232
Conclusion 234
Bibliography 236

List of Figures

Figure 1.1: The Zig zag model of plant immunity 23
Figure 1.2: The Guard, Decoy and Integrated decoy models of NLR activation 28
Figure 1.3: NLR domains found in Plants, Animals and Fungi 30
Figure 1.4: NLR signalling, the known components 38
Figure 1.5: ZAR1 structure and activation 40
Figure 2.1: Vector maps 61
Figure 3.1: Post-infection phenotypes of Col-5 x Ws-2 recombinant inbred lines inoculated with *Albugo candida* isolate AcEM2 75
Figure 3.2: QTL map of *WRR* loci associated with Col-5 x Ws-2 recombinant inbred line resistance phenotypes 76
Figure 3.*3*: Phenotypes of single *WRR* loci containing recombinant inbred lines post infection by *Albugo candida* isolate AcEM2 78
Figure 3.4: *WRR5A* and *WRR5B* genomic clones 80
Figure 3.5: *WRR5A* and *WRR5B* compliment the AcEM2 susceptible phenotype of Ws-2 together but not individually 81
Figure 3.6: Cloning of WRR5A and WRR5B with P-loop and autoimmune mutations 83
Figure 3.7: The P-loops of WRR5A and WRR5B are both required for WRR5A and WRR5B mediated immunity 84
Figure 3.8: Col-*eds1.2-wrr7* lines are susceptible to AcEM2 86
Figure 3.9: *WRR7* Cloning schematic 87
Figure 3.10: *WRR7* is able to complement the susceptible phenotype of Col-*eds1.2-wrr7* plants but not Ws-2 plants 88
Figure 4.1: Overexpression of *WRR5A* and *WRR5B* stimulates an immune response in Ws-2 following *A. candida* infection 98
Figure 4.2: WRR5A and WRR5B form a heterodimeric complex in *Arabidopsis thaliana* 100
Figure 4.3: WRR5B associates with WRR5A-GFP but not RPS4-GFP in *Nicotiana benthamiana* 101

Figure 4.4: WRR5A-YFP and WRR5B-YFP localise to the Plasma membrane in *Nicotiana benthamiana* 104
Figure 4.5: WRR5A and WRR5B localisation in *Arabidopsis thaliana* 105
Figure 4.6: WRR5A and WRR5B require NRG1 to activate defence responses 108
Figure 4.7: CW5 x *adr1 triple* mutant F2 plants show susceptible phenotype to AcEM2 109
Figure 4.8: LC-MS results reveal proteins that are unique to WRR5A-V5 and WRR5B-HF IPs 111
Figure 4.9: Gene ontology analysis of proteins associated with WRR5A and WRR5B by IP-MS 116
Figure 4.10: PIP2 and LOX2 models of hypothesised *WRR5*A and *WRR5B* immune signalling pathways 120
Figure 5.1: *WRR7* expression increases in susceptible and resistant interactions with *Albugo candida* isolates in *Arabidopsis thaliana* 129
Figure 5.2: Principle component and Hierarchical cluster analysis show that RNA seq replicates at each treatment and timepoint are highly similar 132
Figure 5.3: Number of differentially expressed genes after infection by *Albugo candida* 133
Figure 5.4: K-means cluster analysis of RNA sequence data from Col-*eds1.2-wrr7* plants identifies 10 clusters of similarly expressed genes following infection with *Albugo candida* isolate AcEM2 134
Figure 5.5: CAMTA transcription factor binding motifs 137
Figure 5.6: Identification of *Arabidopsis thaliana* Col-*eds1.2* plants susceptible to *Albugo candida* isolate AcEM2 after EMS mutagenesis 140
Figure 5.7: Non-synonymous mutation locations in the EMS candidate proteins 147
Figure 5.8: Cloning of EMS mutant candidate genes 149
Figure 5.9: Complementation of susceptible EMS mutant lines with genomic *CAMTA2* 150
Figure 5.10: Complementation of susceptible EMS mutant lines with genomic *MAP3Kδ4,CHR4* and *MAC7* 151
Figure 5.11: *CAMTA2* is required for *WRR7* immunity in *Arabidopsis thaliana* 155

Figure 5.12: Expression analysis of EMS mutant lines reveals reduced expression of *WRR7* in *camta2*, *chr4* and *mac7* mutants..... 157
Figure 5.13: Cloning and epitope tagging of CAMTA transcription factors and Calmodulin proteins..... 160
Figure 5.14: CAMTA2 and CAMTA3 associate with Calmodulin 2 and Calmodulin 3 161
Figure 5.15: Cloning of *WRR7* 164
Figure 5.16: The C442Y mutation in WRR7 induces autoimmunity and NRG1 is not required for *WRR7* mediated immunity 165
Figure 5.17: Cloning and epitope tagging of WRR7 mutants 167
Figure 5.18: WRR7 localises to the plasma membrane 168
Figure 5.19: The S8F mutation in WRR7 inhibits the cell death induced by the WRR7 autoimmune mutant WRR7-C442Y..... 169
Figure 5.20: MAP3Kδ4 interacts with multiple effector proteins from *Ralstonia psuudosolanacearum* 174
Figure 5.21: Model of *WRR7* mediated immune response to *Albugo candida*.. 176
Figure 6.1: The *Arabidopsis thaliana* DA1 family proteins..... 181
Figure 6.2: WRR7 and WRR5B are part of Clade II of the DA1 family 183
Figure 6.3: Ws-2 lines overexpressing *DA1* family genes display phenotypic abnormalities..... 185
Figure 6.4: Characterisation of Ws-2 lines overexpressing *DAR6:HF*..... 186
Figure 6.5: Immunoblots of DA1 family proteins following *Albudo candida* infection 188
Figure 6.6: Alignment of Peptidase active site motif..... 191
Figure 6.7: Brassicaceae Clade II DA1 family members are evolving faster than Clade I family members..... 192
Figure 6.8: DA1 LIM-Peptidase domain swaps with DAR1, WRR5B and WRR7 194
Figure 6.9: LIM-Peptidase domains of WRR5B and WRR7 are unable to cleave EOD1 195
Figure 6.10: Brassicaceae resistance proteins containing integrated LIM or Peptidase domains..... 197
Figure 6.11: Non-Brassicaceae Resistance proteins contain integrated LIM-peptidase domains 198

Figure 6.12: DA1 family Clade II evolved at the base of the Brassicaceae lineage 201
Figure 6.13: The *Camelina sativa* genome has a copy of *WRR5B* but not *WRR7* 203
Figure 6.14: Phylogeny of Brassicaceae species with resistance genes encoding LIM-Peptidase domains..... 204
Figure 6.15 : Identification of ancestral *WRR5* locus in *Eutrema salsugineum* 205
Figure 6.16: Synteny analysis of *Arabidopsis thaliana* and *Carica papaya* NRG loci 207
Figure 6.17: Domain swaps of the WRR7 LIM-peptidase integrated domain 209
Figure 6.18: Domain swaps of DA1 family LIM-Peptidase domains onto WRR7 N-terminal region induce autoimmunity but not resistance to *Albugo candida.* 210

List of Tables

Table 2.1: List of Primers 62
Table 4.1: AcEM2 susceptible F_2 plants from CW5 x Col-nrg triple mutant and CW5 x Col-adr1 triple mutant crosses are identified 109
Table 4.2: IP-MS V5 IPs identifies novel proteins associated with WRR5A…112
Table 4.3: IP-MS HF IPs identifies novel proteins associated with WRR5B…113
Table 4.4: IP-MS HF IPs identifies novel proteins associated with both WRR5A and WRR5B……114
Table 4.5: IP-MS V5 IPs identifies novel proteins associated with both WRR5A and WRR5B……115
Table 5.1: Transcription factor enrichment analysis reveals that CAMTA transcription factors control the upregulation of genes following *Albugo candida* infection…….. 135
Table 5.2:Bulk segregant analysis by direct sequencing reveals candidate mutant genes responsible for the susceptible phenotypes in each of our *Alblugo candida* isolate AcEM2 susceptible *Col-eds1.2* mutant lines 143
Table 5.3: Complementation of Col-*eds1.2* EMS mutant lines with their candidate genes..... 152
Table 5.4: Identification of susceptible seedlings in F_2 Populations derived from CW14 crossed with T-DNA knock out lines of *camta1/2/3* 154
Table 5.5: Calmodulin 2 is the only calmodulin gene to change in expression in *Arabidopsis thaliana Col-eds1.2-wrr7* plants after infection by *Albugo candida* isolate AcEM2 159
Table 5.6: ADR1 and NRG proteins are not required for *WRR7* mediated immunity 162
Table 6.1: Summary of DA1 family Immunoblot results after infection with *Albugo candida* 189

Acknowledgements

There are many people that have helped me during the course of this PhD. First and foremost, I would like to thank my primary supervisor Dr Volkan Cevik for his continued advice and support throughout this PhD. I would also like to thank the late Prof Rod Scott and Prof Richard Cooper for their support and mentorship at the start of this PhD. In addition, I would like to thank colleagues and friends from the Department of Biology and Biochemistry at the University of Bath for their advice.

The funding for this PhD was provided by the University of Bath, without which this project would not have been possible. Furthermore, I would like to thank the Milner Centre for Evolution for use their facilities during the course of this PhD.

In addition, I would like to thank collaborators on this project, Professor Eric Holub who was involved in setting up this project and performing the initial QTL analysis with Dr Volkan Cevik. I would also like to thank Professor Ksenia Krasileva and Erin Baggs who helped with performing bioinformatic analysis, as well as Nicola Coyle for bioinformatic support and Ellie Jackson-Smith for help in editing the final thesis.

I would also like to thank all the undergraduate and masters students who have been involved in this project, in particular Raphael Ofoe who was always a bubbly presence in the lab, even though he had a tendency to sing the crazy frog song too much!

Finally, I would like to thank all the lab members from the 4S plant lab and the Milner centre plant lab as well as the technical support staff for their support and help throughout this project.

Abstract

Plants utilise receptor proteins to sense invading pathogens and upregulate defence responses. One group of receptors, consisting predominantly of Nucleotide binding-leucine rich repeat receptors (NLRs), acts to recognise the intracellular presence of pathogen derived effector proteins. NLRs consist of a three-domain architecture, comprised of an N-terminal signalling domain, a central nucleotide binding domain that contains a P-loop motif that binds ADP/ATP and acts as an on/off switch for the receptor as well as a C-terminal Leucine rich repeat. The *Arabidopsis thaliana* accession (Col-0) is fully resistant to the biotrophic Oomycete pathogen *Albugo candida*, however other accessions e.g. Ws-2 are susceptible. Recombinant inbred lines (RILs) derived from a cross of these two accessions exhibited multiple different resistance phenotypes. We determined which genes were responsible for the resistance phenotypes and then investigated the mechanisms that the identified genes employ to confer resistance to *A. candida*. We identified three novel *White Rust Resistance* (*WRR*) genes that cause the resistance phenotypes to *A. candida* in Col-0. These include, *WRR5A, WRR5B* and *WRR7*, two of which (*WRR5B* and *WRR7*) encode integrated LIN-11, Isl1 and MEC-3 (LIM)- Zinc-metallopeptidase (Peptidase) domains (LIM-Peptidase). We report that WRR5A and WRR5B form a heterodimeric complex that localises to the plasma membrane in *A. thaliana* and operate by the sensor-helper model of NLR activation. However, unlike other sensor-helper NLR systems the intact P-loops of both proteins are required to stimulate an immune response. In addition, we found that WRR7 stimulates an immune response independent of other NLRs and that after *A. candida* infection this gene is activated by Calmodulin binding transcriptional activator 2 (CAMTA2), independently of CAMTA1 and CAMTA3 which have previously been shown to have functional redundancy with CAMTA2. Furthermore, we implicated Chromatin remodelling protein 4, a Mitogen activated protein kinase (MAP3Kδ4) and MOS 4 associated complex 7 as being involved with the WRR7 resistance mechanism. Therefore, we have shown that the Col-0 genome harbours multiple resistance genes that operate by three distinct mechanisms to cause immunity to one phytopathogen. Expanding our repertoire of distinct resistance gene systems will enable us to understand how NLRs cause immunity to a plethora of different plant pathogens.

This knowledge will pave the way for NLR based engineering approaches to generate novel resistant crop lines in the future.

Abbreviations

NLR: NB-ARC Leucine Rich Repeat Receptor
TNL: TIR containing NLR
CNL: CC containing NLR
CC_R: RPW8 domain containing CNL
TIR: Toll interleukin receptor
CC: Coiled coil
NB-ARC: Nucleotide binding-Apaf resistance to CED4
LRR: Leucine rich repeat
LIM: LIN-11, Isl1 and MEC-3 domain
RPW8: Resistance to powdery mildew 8
PYR: Pyrin domain
CARD: Caspase recruitment domain
NAUGHT: NAIP, CIITA, HET-E and TEP-1 domain
STAND: Signal-transducing ATPase with multiple domains
ID: Integrated domain
ATP: Adenosine triosephosphate
ADP: Adenosine diphosphate
NADP: Nicotinamide adenine dinucleotide phosphate
DA1: Big (Mandarin)
DAR: DA1 related
EOD1/BB: Enhancer of DA1/ Big Brother
MAMP: Microbe Associated Molecular Pattern
MTI: MAMP Triggered Immunity
ETI: Effector Triggered Immunity
ETS: Effector Triggered susceptibility
EBTS: Effector bodyguard Triggered susceptibility
PRR: Pattern recognition receptor
RLK: receptor like kinase
RLP: Receptor like protein
PCD: Programmed Cell Death
HR: Hypersensitive response
BAK1: BRI associated kinase 1
BIR: BAK1 interacting receptor
CERK1: Chitin elicitor receptor kinase
BIK1: Botrytis induced kinase 1
RBOHD: Respiratory burst oxidase homolog D
H3K4me3: histone 3 lysine 4 trimethylation
H3K27me3: histone 3 lysine 27 trimethylation
AcEM2: *Albugo candida* East Mauling isolate 2
WRR: White rust resistance
CW: Col-5 x Ws-2 RIL
SNC1: Suppressor of npr1-1
MOS: Modifier of SNC1
ZAR1: HOPZ activated disease resistance 1
PBL2: PBS1-like protein 2
NRC: NLRs required for cell death

RPS: Resistance to *Pseudomonas syringae*
RRS: Resistance to *Ralstonia solanacearum*
RGA: Resistance gene Analogue
EDS1 Enhanced Disease Susceptibility 1
PAD4: Phytoalexin deficient 4
SAG101**:** Senescence associated gene 101
NDR1: Non-Race-Specific Disease Resistance 1
NRG: N required gene
ADR: Accelerated disease resistance
RPP: Recognition of *Peronospora parasitica*
RPM: Resistance to *Pseudomonas syringae pv maculicova*
SARD: Systemic acquired resistance deficient ICS**:** isochorismate synthase
PAL: phenylalanine lyase
LOX2: Lipoxygenase 2
PIP2: Phosphatidylinositol (4,5)-bisphosphate
HSP: Heat shock protein
CAMTA: Calmodulin binding transcriptional activator
CHR: Chromatin remodelling complex
MAC: MOS4 associated complex
MAPK**:** Mitogen activated protein kinase
MAP3Kδ4: Mitogen activated protein kinase kinase kinase δ 4
CaM: Calmodulin
CPK: Ca2+ dependant protein kinase
CBP: Calmodulin binding protein
CNGC: Cyclic Nucleotide-Gated ion Channels
BRM: Brahma and Kismet
ABF: ABA-responsive element binding transcription factor
PP2C: Protein phosphatase 2 C
SA: Salicylic Acid
ABA: Abscisic acid
JA: Jasmonic acid
ROS: Reactive oxygen species
CHS1: Chilling sensitive 1
HMA: Heavy metal associated domain
NOI: Nitrate induced domain
PHD: Plant homeodomain
SOC3: Suppressor of Chilling sensitive 1
WRKY: WRKY motif containing proteins
RATX: Related to ATX1
RILs: Recombinant inbred lines
CR: Chlorotic resistance
GR: Green resistance
NR: Necrotic resistance
Dpi: Days post infection
IP: Immunoprecipitation
QTL: Quantitative trait loci
ML: Maximum Likelihood
IP-MS: Immunoprecipitation-Mass spectrometry
SNP: Single nucleotide polymorphism
DEG: Differentially expressed gene

BC: Backcross
MS: Murashige and Skoog basal medium
PMSF**:** phenylmethylsulfonyl fluoride
YFP: yellow fluorescent protein
HF: Histidine FLAG tag
GFP: Green fluorescent protein

Chapter 1

General Introduction

Global perspective on land usage: Food security, anthropomorphic population growth and climate change

The world population currently stands at 7.7 billion people and is predicted to rise to 9.7 billion by 2050, with the further addition of another 2 billion people between 2050-2100 (United Nations, 2019). After this century long expansion of the human population, demographic predictions anticipate the stabilisation of the global human population at around 11 billion people (United Nations, 2019). The increase in population will not be evenly distributed across the globe, Western Europe and the Americas will experience negative population growth over this time period with most of the population growth occurring in Asia and sub-Saharan Africa (United Nations, 2019). The burgeoning global population causes challenges to global food security due to the expanding demand for food. To meet the demand for future consumption, the production of food has to increase to meet the dietary requirements of the global population. Not only is the world population growing, but it is becoming more affluent, causing a shift in global dietary trends towards a higher calorific diet that consumes more meat products. Dietary shifts create more strain on food production systems because the farming practices needed to generate produce that meets the demands of diets consumed by more affluent people, require greater areas of land compared to lower calorific diets predominantly consisting of arable produce (Myers and Kent, 2003; Alexandratos and Bruinsma, 2012). To meet this demand, it is predicted that per capita food production needs to increase by 15%. Taking into account the population increase, this scenario would result in a 60% increase in global food production by 2050 (Alexandratos and Bruinsma, 2012). Not only is population growth a concern for food security but also land area and land-use conflicts are becoming a more prominent issue. The area of cultivatable land is becoming increasingly smaller as we increase the global agricultural land area to meet the demands of a growing population, and increases in the cultivation of new land is predicted to decline in the coming decades (Alexandratos and Bruinsma, 2012). Increasing the global land area that is farmed is problematic as the remainder of Earths cultivatable land contains some of the planets most important ecosystems

such as tropical forests and wetlands (Gibbs et al., 2010). On top of this, the need to tackle climate change is resulting in more and more land being set aside for the production of biofuels and there is an increasing push to re-forest land to tackle the global climate crisis particularly in Europe and China (Searchinger et al., 2008; Song et al., 2018). The combination of these factors means that we need to find new ways to intensify agriculture from our current available land in order to support the worlds growing population without destroying the remainder of Earths undisturbed ecosystems.

One way of increasing global yields, is to produce crop varieties that are more resistant to pests and diseases. It is predicted that up to 30% of global crop yields are lost due to pests and diseases (Savary et al., 2019). Agroecosystems are evolutionary biased in favour of pathogens, due to the abundance of monocultures resulting in host genetic homogeneity. The lack of genetic diversity in crops creates an environment conducive to the evolution of pathogens, as new strains can undergo rapid selective sweeps that devastate crops (Zhan and McDonald, 2013). Therefore, we have to be constantly breeding and developing new crop varieties to combat the rapid evolution of crop pests and diseases. Innovations aimed at reducing the global losses of crops to diseases will help contribute to feeding the worlds growing population and increase global yields by minimising losses caused by phytopathogens. To breed or engineer crop varieties that are resistant to pathogens without reducing yields, requires an intricate knowledge of the plant immune system and the host-pathogen interactions that result in susceptibility or resistance. Therefore, it is imperative that we gain a greater understanding of the molecular interplay that occurs at the plant-pathogen interface to inform crop breeding approaches.

Current knowledge of the plant immune system

Responses to invasion and the classic model of plant immunity

Phytopathogens cause molecular disruption within a host when they invade. The disruption to host tissue and cells results in the release of signalling molecules that stimulate the upregulation of plant defences. Plant signalling events revolve around the release of small ions or molecules such as Ca^{2+}, reactive oxygen species (ROS) or phytohormones such as Salicylic acid (SA) that interact with components of the plant immune system to trigger defence responses (Camejo, Guzmán-Cedeño and Moreno, 2016; Kudla et al., 2018). The recognition of pathogens by plants stimulates several defence responses that operate locally at the site of infection or more globally at the whole plant level. Small signalling molecules such as SA play a role in upregulating both of these responses (Gao, Q.-M. et al., 2015). At the broad level these small signalling molecules result in systemic acquired resistance at the whole plant level by the upregulation of general plant defences such as the production of antimicrobial compounds such as glucosinolates in the crucifers (Singh, Guest and Copeland, 2015). At the local level these signalling molecules upregulate defences in the plant innate immune system that result in strong immune responses that prevent pathogen invasion into the host (Jones, J., D. G. and Dangl, 2006).

Plant hormones play crucial roles in regulating plant developmental and stress responses including the regulation of plant immunity. Two of the most important phytohormones involved in mediating the plant immune system are SA and Jasmonic acid (JA) which govern responses to biotrophic pathogens and necrotrophic pathogens respectively (Glazebrook, 2005). Other plant hormones also play important roles in plant immunity by interacting with SA and JA to modulate the severity of defence responses (Berens et al., 2017). For example, Abscisic acid (ABA) acts antagonistically to SA and can repress immunity triggered by increases in SA (Yasuda et al., 2008; Jiang et al., 2010). Although the role of each hormone can be more nuanced with ABA also important for stomatal responses to pathogen invasion (Lind et al., 2015). Plant growth is supressed during an immune response (Albrecht and Argueso, 2017), therefore plant hormones that are involved in stimulating growth such as auxins and ABA often have negative roles in plant immunity (Berens et al., 2017). Other hormones such as cytokinin's, ethylene, gibberellins and brassinosteroids play conflicting roles that are often species and

condition dependent and whether they have a wider role in plant immunity is still being determined (Berens et al., 2017). However, the interplay between all the phytohormones is critical for the response of a plant to its environment and therefore this network of interactions plays a critical role in determining whether to activate the plant immune system.

Our current understanding of the plant innate immune system was first outlined in the zig-zag model of plant immunity (Fig 1.1). The zig-zag model partitions the molecular interactions between host and pathogen into two distinct layers that operate simultaneously during invasion (Jones, J., D. G. and Dangl, 2006). The first layer occurs at the plasma membrane, where pattern recognition receptors (PRRs) recognise microbe associated molecular patterns (MAMPs) and then stimulate the induction of basal plant defences known as MAMP triggered immunity (MTI), that are typically strong enough to generate resistance against non-specialised pathogens. Some pathogens can overcome this layer of immunity by secreting effector proteins into the cell that interfere with MTI, resulting in Effector triggered susceptibility (ETS). The second layer of plant immunity operates intracellularly. It is controlled predominantly by nucleotide binding-leucine rich repeat receptors (NLRs) or NLR derivatives that recognise the presence of effectors, this layer of immunity is known as Effector triggered immunity (ETI). Since the zig-zag model was first proposed, another step has been added, as pathogens have evolved ‘effector bodyguards’, proteins whose role is to mask the presence of effectors in the cell (Paulus, Kourelis and van der Hoorn, 2017). Consequently, effectors are able to evade detection by NLRs, resulting in Effector bodyguard triggered susceptibility (EBTS).

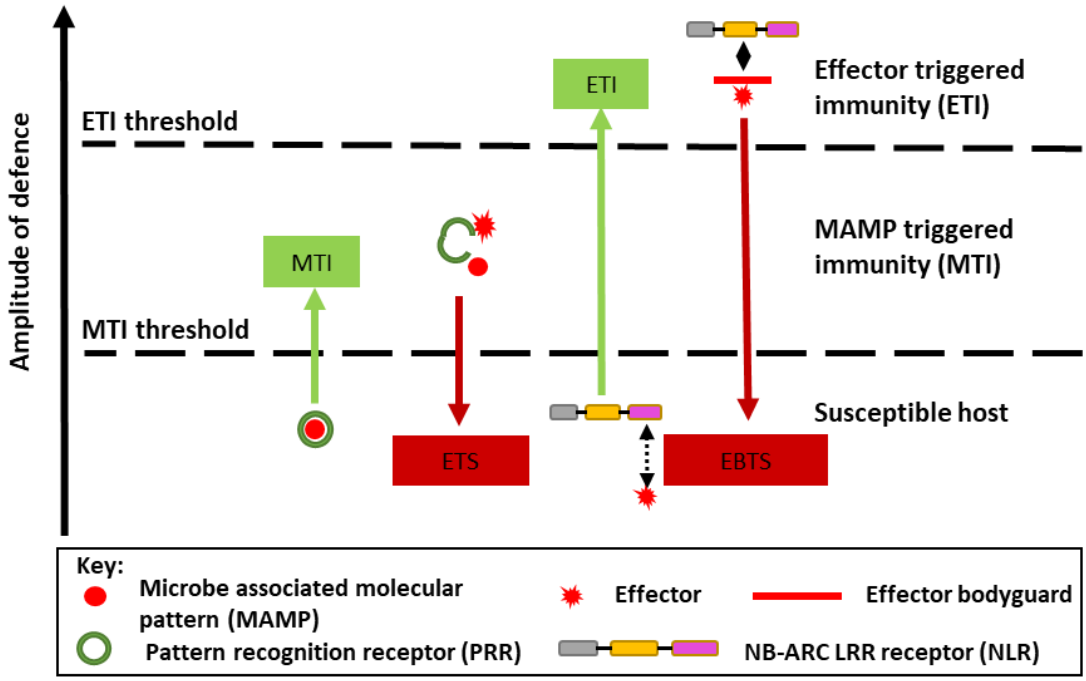


Figure 1.1: The Zig zag model of plant immunity
The Zig Zag model of plant immunity adapted from (Jones, J., D. G. and Dangl, 2006). Green symbols represent plant proteins that are required for MAMP triggered immunity (MTI) or effector triggered immunity (ETI), red markers represent pathogen derived molecules that lead to effector triggered susceptibility (ETS) or effector bodyguard triggered susceptibility (EBTS).

Cell surface receptors and the first layer of plant innate immunity

The initial layer of plant innate immunity is governed by the recognition of MAMPs by a group of cell surface localised receptor like kinases (RLKs) or Receptor like proteins (RLPs) known as PRRs (Zhou, Tang and Wang, 2017). MAMPS are a group of molecules that are indicative of groups of taxa containing plant pathogens, for example bacterial flagellin, fungal chitin and Oomycete elicitins or β-1,3/1,6 glucans (Jones, J., D. G. and Dangl, 2006; Sánchez-Vallet, Mesters and Thomma, 2015; Derevnina et al., 2016; Judelson and Ah-Fong, 2019; Wang, Yan, Tyler and Wang, 2019). Once MAMPS are detected by PRRs a signal cascade is induced leading to the activation of basal plant defences, otherwise known as MTI. MTI elicits defence responses that are strong enough to prevent infection by non-host pathogens. The binding of PRRs to their associated MAMP ligand often involves plasma membrane localised co-receptors, for example the binding of bacterial MAMP flg22 (a peptide of flagellin) to associated PRR FLS2 requires co-receptor BRI1 associated kinase 1 (BAK1) and fungal MAMP chitin requires co-receptor

Chitin elicitor receptor kinase 1 (CERK1) to associate with corresponding LysM motif containing PRRs in order to stimulate downstream defence responses (Gómez-Gómez and Boller, 2000; Couto and Zipfel, 2016; Zipfel and Oldroyd, 2017). Co-receptors such as BAK1 and CERK1 form important signalling hubs and interact with multiple ligand-PRR complexes (Smakowska-Luzan et al., 2018). These hubs are important because they are crucial to this initial layer of non-specific pathogen recognition and can be exploited by pathogens to circumvent MTI during invasion.

Once a microbe is recognised at the cell surface, MTI then induces changes in small signalling molecules such as Ca^{2+} and ROS. Alterations in Ca^{2+} homeostasis are among the first changes caused by the perturbance of plant cells by microbes. After microbial disturbance, Ca^{2+} crosses the plasma membrane through CNGC channels, causing a cytosolic Ca^{2+} influx (Tian, W. et al., 2019). Ca^{2+} is an important second messenger, that regulates several developmental and stress response pathways. The induction of Ca^{2+} into the cell triggers and co-ordinates MTI responses along with co-activators such as Botrytis induced kinase 1 (BIK1) which becomes rapidly phosphorylated after binding to BAK1 and CERK1 (Zhang, J. et al., 2010; Kadota, Shirasu and Zipfel, 2015). One prominent factor in MTI signalling that is regulated by Ca^{2+} is the generation of ROS by pathogen disturbance of the plant cell wall and plasma membrane that are recognised by the Respiratory burst oxidase homolog D (RBOHD) NADP oxidase (Saijo, Loo and Yasuda, 2018). RBOHD is a calcium ion acceptor that binds Ca^{2+} to calcium binding EF-hand domains that it contains at its N-terminus (Kadota, Shirasu and Zipfel, 2015). As well as physically binding Ca^{2+}, RBOHD is phosphorylated by BIK1 and Ca^{2+} dependant protein kinases (CPKs) that control RBOHD activity, thereby modulating the ROS response in a Ca^{2+} dependant manner (Kobayashi et al., 2007; Kadota et al., 2014; Li, L. et al., 2014; Kadota, Shirasu and Zipfel, 2015). These small signalling molecules then go on to stimulate intracellular signalling responses such as Mitogen activated kinase (MAPK) cascades that result in the upregulation of MTI responses through the release of transcription factors such as the WRKY transcription factors (Meng and Zhang, 2013; Saijo, Loo and Yasuda, 2018).

Suppression of MTI by invading pathogens

For a plant pathogen to be able to colonise a host it first must evade or suppress immunity triggered by MTI. To accomplish this, pathogens secrete a plethora of

molecules known as effector proteins, whose role is to disrupt host cell functions, aid infection and promote virulence for the benefit of the invading pathogen (Jones, J., D. G. and Dangl, 2006; Toruño, Stergiopoulos and Coaker, 2016). Effectors are secreted into host cells using specialised structures such as the bacterial type III secretion system or fungal/Oomycete specialised feeding structures known as haustoria (Büttner and He, 2009; Presti et al., 2015; Judelson and Ah-Fong, 2019). Exactly how intracellular fungal and Oomycete effectors are translocated across the extra-haustorial space and across the plants plasma membrane is still and active area of research (Rodriguez-Moreno et al., 2017). Traditionally effectors have been considered to exclusively be proteins (Toruño, Stergiopoulos and Coaker, 2016). However, recent studies have identified siRNAs secreted from fungal necrotrophic pathogen *Botrytis cinereal* that have been found to silence tomato and *A. thaliana* immune associated genes, essentially performing the role of an effector (Plett and Martin, 2017). Therefore, our understanding of how pathogens manipulate their hosts is changing and whether the definition of an effector should be expanded to include non-protein molecules is up for debate. Intracellular effector targets are diverse and include transcription factors, immune receptors and plant hormones, particularly defence related hormones salicylic acid and jasmonic acid (Toruño, Stergiopoulos and Coaker, 2016). These targets are usually part of key processes within the plant including the upregulation of immune responses or transportation of nutritional molecules which the pathogen is attempting to obtain (Chen, L.-Q. et al., 2010; Plett and Martin, 2017; Walerowski et al., 2018). One of the most important functions of effectors, is to suppress MTI responses that have been triggered during the colonisation of the host tissue. There are many well characterised examples of MTI suppression by effectors and several key MTI signalling nodes have been identified which pathogens target such as the BAK1-FLS2 and BIK1 complex which is targeted by bacterial pathogens *Pseudomonas syringae*, *Xanthomonas oryzae* and *Xanthomonas campestris pv campestris* (Göhre et al., 2008; Lu et al., 2010; Macho et al., 2014; Wang, G. et al., 2015; Li, L. et al., 2016; Üstün et al., 2016; Irieda et al., 2019). As well as suppressing MTI, pathogens have evolved to evade recognition by PRRs, for example *Magnaporthe oryzae* secrets effector Secreted LysM Protein1 which sequesters the MAMP chitin, preventing its recognition by the chitin triggered PRRs (Mentlak et al., 2012). As a consequence of their function in evading immune signalling, effector proteins are highly important to pathogenic infection strategies

and the loss of effector functions can be detrimental to the ability of a pathogen to invade a host. However, effectors operate within host tissues and are regularly interacting with host proteins. Therefore, as well as being vital for the virulence of a pathogen they represent an Achilles heel that can be exploited by the host, providing another interaction interface where a pathogen can be recognised. This interaction has resulted in the evolution of a second layer of plant immunity.

The second layer of plant immunity: detecting the effector

The second layer of plant immunity centres around the detection of pathogen effector proteins by intracellular Nucleotide binding leucine rich repeat receptors (NLRs), previously known as nod like receptors (Ting et al., 2008). NLRs are a large conglomerate of proteins which act as an intracellular detection network that recognise invading pathogens and activate immune responses through an increase in the phytohormone Salicylic acid (SA) (Vlot, Dempsey and Klessig, 2009). Immune responses that are triggered by the recognition of an effector by an NLR lead to rapid 'non autolytic' programmed cell death of infected cells, otherwise known as the hypersensitive response in plants, which kills the invading pathogen (van Doorn, 2011). Due to this highly specific recognition of pathogen effectors by the plants intracellular immune system, this 'layer' of plant immunity has come to be known as effector triggered immunity (ETI).

NLRs sense the presence of pathogen effectors by several different mechanisms (Fig 1.2), either by directly interacting with the effector molecule or by indirectly perceiving the cellular effects of an effectors action (Baggs, Dagdas and Krasileva, 2017). The majority of these interactions act through indirect association of the NLR with their cognate effector. The two most prevalent mechanisms by which an NLR can indirectly sense a pathogens presence by the recognition of an effector, is through the guard or decoy models of NLR activation (Jones, J., D. G. and Dangl, 2006; Van Der Hoorn and Kamoun, 2008; Jones, J.D.G., Vance and Dangl, 2016). These models describe the association of an NLR with an effector through an intermediary protein called a guardee or a decoy (Fig 1.2). Guardee proteins are host proteins that are targeted by plant pathogens during infection and are often part of the basal plant immune system or involved in the processing of nutrients through the plant. Whilst guardee proteins are monitored by NLRs they retain their normal host function. On the other hand, decoy proteins are thought to be proteins that mimic a

host guardee but have lost their original function (Fig 1.2). This allows a host to control the level of decoy protein in the cell without interfering with any of its developmental programmes. The effector proteins will then not only have a strong binding affinity to their intended target protein but will also bind to the structurally similar decoy protein, activating its associated NLR leading to ETI. Sometimes, the decoy or decoy domains have evolved to become part of the NLR architecture acting as an integrated decoy or bait domain for the NLR, resulting in a third model of NLR activation called the integrated decoy model (Cesari et al., 2014). The integrated decoy model of NLR activation is mainly thought to be triggered by direct interaction of the effector with the integrated domain (ID) although there are emerging examples where ID containing NLRs can also be involved in indirect association between NLR and effector (Fujisaki et al., 2017). The integrated domain needs to be a decoy domain that has been divested of its original host function as the domain will now be under the regulatory control of the immune system and not its original regulatory mechanism. If the domain retained its original function it could negatively impact the plants development and regulatory functions, therefore it is unlikely we would see the evolution of an integrated-guardee domain.

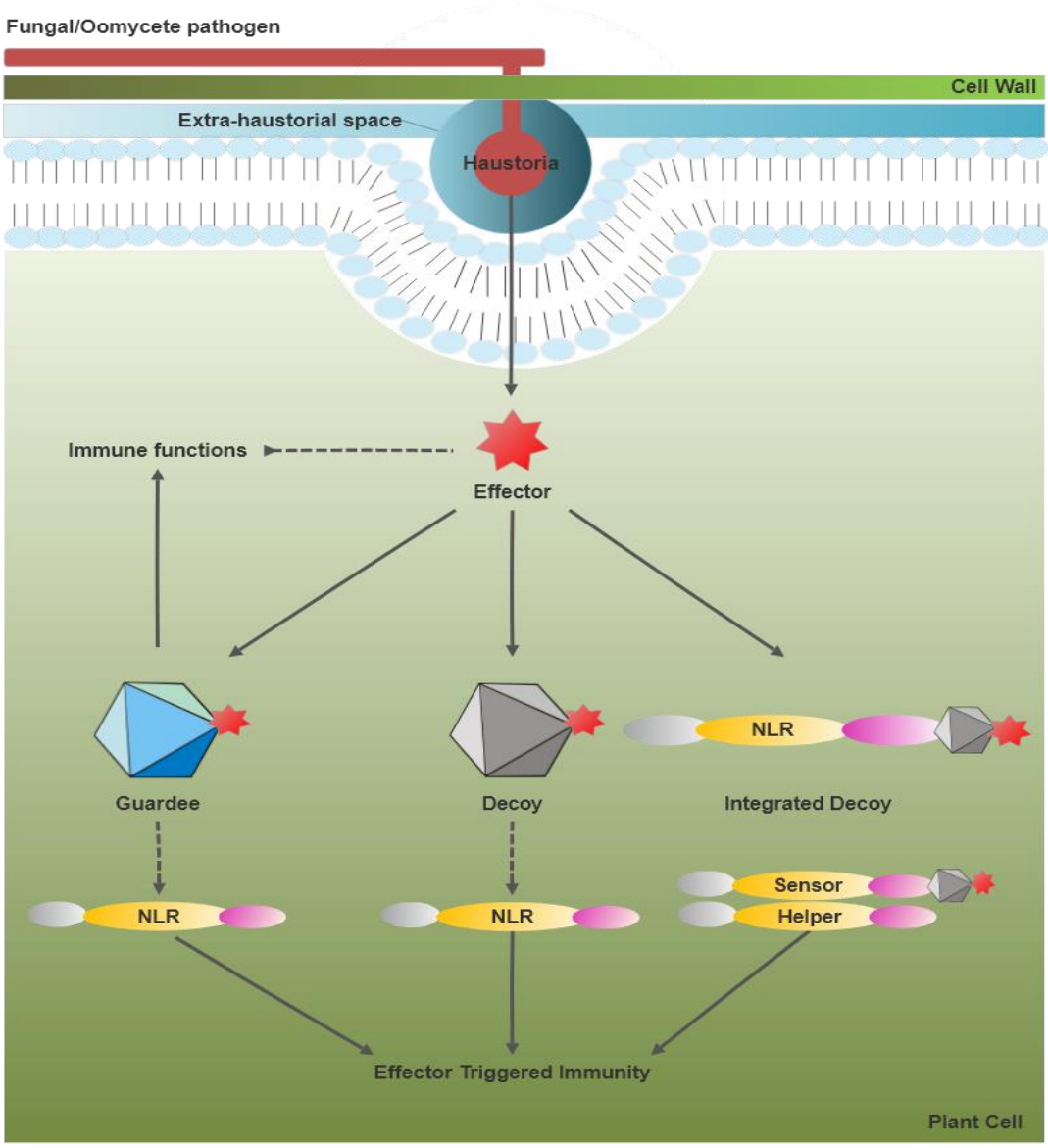


Figure 1.2: The Guard, Decoy and Integrated decoy models of NLR activation Model showing the different mechanisms that NLRs employ to recognise the presence of fungal or Oomycete effector proteins (red star) in the cell. Solid arrows show direct interactions, dashed arrows show were interactions can be indirect. Effectors are secreted from the haustoria and are translocated intracellularly. The effector then binds to its target proteins that predominantly has a role in basal plant immunity or MAMP triggered immunity. If the effector target protein is monitored by the NLR either directly or indirectly then this protein is known as a guardee and its associated NLR activates effector triggered immunity (ETI) this is known as the guard model. In the decoy model, the effector can't distinguish between its host target and a structurally similar protein known as a decoy and the interaction of the decoy protein and effector is recognised by an associated NLR and ETI is activated. In the integrated decoy model, the decoy protein or part of a decoy protein has become integrated into the structure of an NLR and the effector interacts with this domain. This NLR can then either directly stimulate immunity or require a helper NLR to activate ETI. Model adapted from (Van Der Hoorn and Kamoun, 2008; Cesari et al., 2014).

NLRs throughout the Kingdoms

NLRs are not exclusive to plants but are found in many different organisms including animals and fungi (Jones, J.D.G., Vance and Dangl, 2016; Meunier and Broz, 2017; Uehling, Deveau and Paoletti, 2017). They have a distinct canonical structure containing a central nucleotide-binding domain, an N-terminal signalling domain and a C terminal Leucine rich repeat (LRR). Although the basic NLR architecture is similar between plants, animals and fungi, distinct differences have been identified in their constituent domains (Fig 1.3). The N-terminal domain of NLRs, performs an active role in signalling downstream defences and these signalling domains show the greatest divergence between the kingdoms (Fig 1.3). In the Plantae kingdom, NLRs have two predominant N-terminal domains, the Toll interleukin receptor (TIR) or the coiled-coil (CC) domain, that enact downstream signalling events. In animals, the equivalent N-terminal signalling domain to the CC/TIR domains found in plants are the pyrin (PYR)/caspase recruitment domain (CARD) (Jones, J.D.G., Vance and Dangl, 2016). Whereas in fungi the N-terminal signalling domains are found in a much greater diversity, with at least twelve domains identified, some containing similarity to the TIR and CC domains of plants (the Het and HeLo domains respectively), while other domains such as the phospholipase patatin domain are completely unique at this position in NLRs in fungi (Dyrka et al., 2014; Uehling, Deveau and Paoletti, 2017). This divergence in signalling domains means that the downstream responses triggered by NLRs in the different kingdoms occur through evolutionary distinct mechanisms.

NLRs not only differ at their N-termini between the kingdoms but also at their nucleotide-binding domain (Fig 1.3). The central nucleotide binding domain of both plants and animal NLRs are associated with ADP in their ‘off’ state or ATP in their ‘on’ state (Hu, Z. et al., 2013; Wang, J., Wang, et al., 2019). The association of NLRs with ADP/ATP makes them part of a larger class of proteins known as signal transducing ATPase with multiple domains (STAND) proteins (Bentham, A. et al., 2016). In plants, the central nucleotide-binding domain is the NB-ARC (nucleotide binging-Apaf1 resistance to CED4) domain which contains a phosphate-binding loop (P-loop) and Walker B motif. In animals, the STAND domain is represented by the NACHT domain (named after its presence in NAIP, CIITA, HET-E, TP1 proteins), rather than the NB-ARC domain found in plants whereas fungi genomes harbour

NLRs with both the NACHT and NB-ARC domains (Dyrka et al., 2014; Jones, J.D.G., Vance and Dangl, 2016).

Common to NLRs between the kingdoms is the presence of a highly repetitive region at their C-terminus either a LRR domain in animals and plants or a WD, ANK or TPR domain in fungi (Dyrka et al., 2014). Structural analyses have revealed that the C-terminus of NLRs mainly plays a role in maintaining the NLR in an autoinhibited monomeric state or provides an interaction site between NLRs and their interactors in both plants and animals (Hu, Z. et al., 2013; Jones, J.D.G., Vance and Dangl, 2016; Wang, J., Wang, et al., 2019).

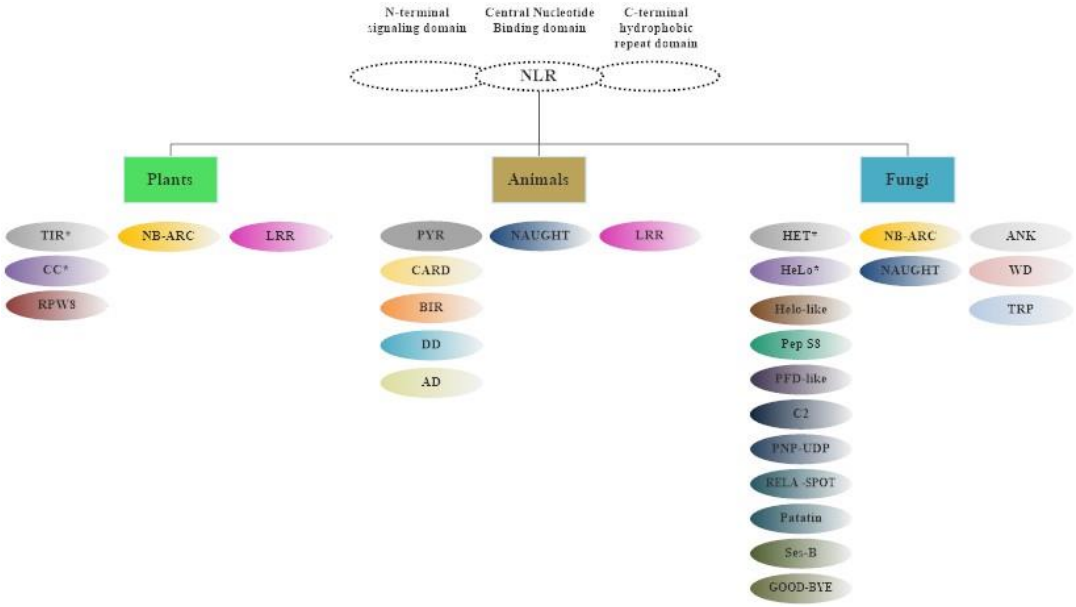


Figure 1.3: NLR domains found in Plants, Animals and Fungi
Diagram showing the diversity of protein domains found at the N-terminal, central and C-terminal positions of NLRs in the Plantae, Animalia and Fungi kingdoms. Homologous domains are shown by colouration and domains showing close homology between the kingdoms are marked with an *. Figure adapted from (Uehling, Deveau and Paoletti, 2017).

On the Origin of NLRs

The origin of NLRs across the kingdoms is not yet fully resolved. All constituent parts of NLRs can be found in lineages that predate the divergence of prokaryotes and eukaryotes although the fusion of these domains into NLRs occurred later in the eukaryote lineages (Yue et al., 2012). The fact that differing NB domains are present in different kingdoms suggests that NLRs evolved convergently between these lineages and although they are structurally similar, caution should be taken in drawing comparisons between these systems (Urbach and Ausubel, 2017). There is still debate on the origin of plant NLRs (Gao, Y. et al., 2018; Shao et al., 2018), with the latest studies suggesting that they emerged in a common ancestor predating the origin of green plants (Shao et al., 2018). However, despite the lack of resolution on the evolutionary origins of plant NLRs there is lots to be learnt about host-pathogen co-evolution from the combination of studying the evolution of NLRs across the kingdoms of life.

In plants, NLRs are divided into distinct groups depending on the presence or absence of the TIR domain at their N-termini (Fig 1.4). Plant NLRs containing a TIR domain are collectively termed TIR-domain containing NLRs (TNLs), the remainder of plant NLRs have no N-terminal TIR domain instead the majority of these contain a CC domain, therefore non-TNL NLRs are often referred to as CC-containing NLRs (CNLs). A third sub-group of NLRs that are a sister group to the CNLs, have been identified more recently, this group contain a resistance to powdery mildew 8 (RPW8) domain at their N-termini and are called RPW8 containing CNLs or CNLs containing an RPW8 domain (RNLs/CC_Rs) (Shao et al., 2014; Shao et al., 2016; Zhang, Y.-M. et al., 2016; Nepal et al., 2017). The evolution of these three NLR groups in plants has occurred independently (Shao et al., 2018). TNLs have to date only been identified in dicots, whereas CNLs and CC_Rs are present in monocots as well as dicots. It is therefore logical to draw the conclusion that non-TNLs predate the evolution of TNLs. However, this does not seem to be the case, as TNLs show closer similarity in their NB-ARC domain sequences to basal land plants, suggesting that the TNLs are a more ancient group of NLRs that have subsequently been lost in the monocot lineage (Yue et al., 2012). The distinction of these three groups of NLRs (TNLs, CNLs and CC_Rs) is important because all three groups signal through different pathways to stimulate a SA response leading to cell death (Fig 1.4). We

know most about the downstream signalling network employed by TNLs whereas the signalling networks utilised by CNLs and CC_R NLRs are less well understood.

NLR encoding genes are an ancient group of defence genes. However, they have diverged massively since their origin, generating many variants and adapting new functions in recognising pathogens. In order to achieve this, they have to have evolutionary plasticity in order to combat rapidly evolving pathogens. Therefore, NLRs themselves are prone to rapid evolution and are highly polymorphic leading to the generation of several allelic variants within populations (Van de Weyer et al., 2019). This innate plasticity and the ability of NLRs to evolve swiftly has provided many different variations of NLRs that don't always conform to the canonical domain structure that we see most predominantly in plants. However, these NLR derivatives are often still crucial components of plant defence.

NLR-like proteins and the diversity of NLR architectures

The vast majority of NLRs are formed of the canonical three domain structure, some of which contain an extra integrated domain as discussed earlier. However, there are NLR-like genes within plant genomes that contain some but not all of the typical NLR domains. In the *A. thaliana* pan-NLRome alone there are 97 different known NLR or NLR-like architectures (Van de Weyer et al., 2019). Some of the NLR-like proteins still retain the ability to activate defence responses like their NLR counterparts (Nandety et al., 2013). The most prevalent NLR-like architectures are TNs or CNs, NLRs lacking an LRR domain. These proteins occur in relatively small amounts compared to full NLRs but they still form a substantial group of resistance proteins with over 1257 TNs identified in the *A. thaliana* pan-NLRome alone (Van de Weyer et al., 2019). One of the roles of NLR-like proteins is to act in concert with full length NLRs to cause resistance, for example TN protein chilling sensitive 1 (CHS1) interacts with TNL suppressor of *chs1-2* (also known as SUSA or WRR12) to bring about an immune response indicated by upregulation of SA and pathogenesis related genes (Wang, Yuancong et al., 2013; Zbierzak et al., 2013; Zhang, Y. et al., 2016). One TN protein, TN2 has been shown to monitor the homeostatic state of calcium during infection through its interaction with CPK5 and activate ETI responses (Liu et al., 2017). The study of these unusual resistance proteins offers an insight into the functionality of each structural element within

NLR architectures and provides insights into the role of individual domains during NLR activation.

Immune signalling events in phyto-NLR systems

Signalling networks are highly important in governing developmental and stress response programmes in eukaryotic organisms. In plants, signalling networks triggered by an NLR detecting an invading pathogen result in cell death and are therefore tightly regulated. Salicylic acid signalling networks are critical for the activation of cell death stimulated by NLRs. However, the signalling networks leading to SA accumulation are different depending on the NLR N-terminal domain (Fig 1.4). The identification of components involved in the mechanistic action of CNL, TNL and CC_R signalling networks is still a highly active area of research and understanding the signalling networks and mechanisms that plant NLRs utilise to bring about defence activation is important for future crop engineering approaches.

TNL downstream responses are obligate on lipase like protein Enhanced disease susceptibility 1 (EDS1) and its formation of mutually exclusive heterodimeric complexes with either Phytoalexin deficient 4 (PAD4) or PAD4-related senescence-associated gene 101 (SAG101) (Wagner et al., 2013; Lapin et al., 2019). Intriguingly, two families of CC_Rs, the N gene required (NRGs) and Accelerated disease resistance (ADR1s) have also been implicated in TNL downstream signalling (Fig 1.4) and are involved with TNL defence responses stimulated by EDS1-PAD4 and EDS1-SAG101 complexes (Castel et al., 2018; Wu, Z. et al., 2018). In the Solanaceae, EDS1-SAG101 complexes are active downstream of TNL immune signalling and require NRGs for immune signalling whereas EDS1-PAD4 complexes are reliant on ADRs (Gantner et al., 2019; Lapin et al., 2019). Both the NRG and ADR1 gene families have thus been implicated in TNL signalling. On top of this, NRGs are absent from plant families containing TNLs, in contrast to ADR1 family proteins, that are present in non-TNL containing plant families, suggesting that the function of ADR1 family proteins in plant disease resistance may be broader than what is currently documented (Wu, Z. et al., 2018). It is also important to note that the Brassicaceae family PAD4 is quite divergent from Solanaceous PAD4 and in Solanaceae species the EDS1-SAG101 complex is required for TNL signalling, whereas Brassicaceae species require the EDS1-PAD4 complex for TNL signalling (Wagner et al., 2013; Gantner et al., 2019). Therefore, a divergent signalling

pathway has evolved in the Brassicaceae lineage compared to the Solanaceae (Gantner et al., 2019). This divergence in immune signalling in the Brassicaceae which includes the most studied plant model species *A. thaliana* means that any finding relating to TNL signalling in this lineage may not be easily applicable across the family boundary. More recently, TIR domains have been found to have NAD^+ catalytic activity in both animals and plants (Horsefield et al., 2019; Wan et al., 2019). Plant TIR domains of TNLs Resistance to *Pseudomonas syringae* 4 (RPS4) and Recognition of *Peronospora parasitica* 1 (RPP1), have been shown to deplete NAD^+ in *E. coli* and produce breakdown products nicotinamide and cyclic ADP-ribose a finding that was replicated with TIR domains from Flax and grapevine NLRs L6 and RUN1 (Horsefield et al., 2019; Wan et al., 2019). On top of this, the breakdown product of this reaction cyclic ADP-ribose can trigger Ca^{2+} influxes, providing a potential signalling mechanism that could lead to the induction of defence related responses triggered by TNLs (Hunt, Lerner and Ziegler, 2004; Horsefield et al., 2019).

The signalling of CNLs is less well understood with a small amount of signalling partners identified that are only required for a subset of CNLs to trigger defence responses (Fig 1.4). An example of this is plasma membrane anchored protein, Non-race specific disease resistance 1 (NDR1) which is required for the downstream signalling of two CNLs Resistance to *Pseudomonas syringae* 2 (RPS2) and Resistance to *Pseudomonas syringae pv maculicova* 1 (RPM1) but this is not a general requirement of other CNLs (Day, Dahlbeck and Staskawicz, 2006b; Knepper, Savory and Day, 2011; Cui, Tsuda and Parker, 2015). Even less is known about the downstream signalling of CNL sister group CC_Rs other than the involvement of the NRGs and ADR families in TNL downstream signalling and even then, the mechanism that they utilise in TNL signalling is not yet known.

Although TNLs and CNLs signal through distinct pathways, there are some common downstream signalling components between the two systems. For example, both systems utilise MAPK cascades to activate defence responses. In plants, MAPK cascades involve a three tiered response, whereby MAPKs are activated by MAPKKs that are in turn activated by MAPKKKs that are activated by an external stimulus (Zhang, S. and Klessig, 2001). MAPK cascades are a crucial signalling hub in plants that mediate responses to various environmental stimuli and are integral to

the plant immune system in responding to both MTI and ETI stimuli (Bi, G. and Zhou, 2017; Wang, W. et al., 2020). There is substantial cross talk between these two immune hubs for example two MAPKs, MPK3 and MPK6 are active during both MTI and ETI and only the duration of their activation determines the type of response, with longer activation leading to an induction of salicylic acid and ETI responses whereas shorter activation periods result in MTI responses (Tsuda et al., 2013). Specific MAPKs have been implicated in direct downstream signalling responses in ETI e.g. the MAPKKKα is important for the immune function of tomato CNL PRF and MPK3 and MPK6 are both important for RPS2 mediated immunity in *A. thaliana* (del Pozo, Pedley and Martin, 2004; Tsuda et al., 2013; Peng, Yujun, van Wersch and Zhang, 2017). However, exactly how these signalling cascades link CNL and TNL mediated immunity with downstream salicylic acid responses is still not fully understood.

Both CNLs and TNLs activate cell death responses through a spike in phytohormone SA (Fig 1.4), the link between TNLs, CNLs and the increase in SA is still not resolved (Loake and Grant, 2007). In plants, SA can be synthesised by two pathways in the chloroplast (the ICS or PAL pathways), by conversion of isochorismate to SA by isochorismate synthase (ICS) and another unknown enzyme which is thought to be similar to isochorismate pyruvate lyase found in bacteria (Yamasaki et al., 2013) or by conversion of cinnamate produced from phenylalanine lyase activity (PAL) (Chen, Z. et al., 2009). The primary SA biosynthesis pathway in plants is believed to be the ICS pathway, in *A. thaliana* the ICS pathway is believed to account for >90% of free SA, but in Soybean the production of SA is thought to be equally mediated between the two pathways, and experiments in both *A. thaliana* and *N. benthamiana* have shown that the PAL pathway is still important in the production of SA during defence responses (Wildermuth, M., C. et al., 2001; Chen, Z. et al., 2009; Shine et al., 2016). Once SA is produced, it is transported across the chloroplast envelope via the EDS5 transporter into the cytoplasm where a range of ETI stimulated defence responses are initiated (Yamasaki et al., 2013). Key genes involved in SA biogenesis, including *ICS1* and *EDS5* are regulated by two transcription factors, Systemic acquired resistance deficient 1 (SARD1) and Calmodulin binding protein-60-like-g (CBP60g), both of these transcription factors are from the Calmodulin Binding protein family which are responsive to changes in cellular Ca^{2+} levels (Ding

and Redkar, 2018). Therefore, the Ca^{2+} signalling network is a key network that could provide the link between NLRs, SA and downstream signalling events (Cheval et al., 2013).

Ca^{2+} is a crucial second messenger molecule in plants that has diverse functions, particularly in co-ordinating responses to external stimuli (Kudla et al., 2018). Biotic interactions cause changes in the cellular Ca^{2+} levels, for example it is well established that nuclear Ca^{2+} spiking events are crucial for the development of symbiotic plant-microbe interactions (Oldroyd, 2013). Fluctuations in Ca^{2+} levels have not only been recorded in beneficial plant microbe interactions but also in pathogenic interactions, for example Ca^{2+} undergoes a cytosolic influx through plasma membrane calcium channels such as the CNGC2 and CNGC4 channel in *A. thaliana* during MTI (Yuan et al., 2017; Hander et al., 2019; Tian, W. et al., 2019). CNGC Ca^{2+} channels have also been shown to be regulated by BAK1 and BIK1, key co-regulators of MTI signalling (Meena et al., 2019; Yu, X. et al., 2019). MTI induced Ca^{2+} signals can be amplified in a positive feedback loop with the release of ROS during invasion through the regulation of Ca^{2+} controlled NADP oxidase RBOHD (Kadota, Shirasu and Zipfel, 2015; Zipfel and Oldroyd, 2017). The release of Ca^{2+} in MTI activates downstream signalling, particularly during a herbivory response where type-II metacaspases are activated by Ca^{2+} signals and subsequently cleave Pep signals from precursor proteins that then interact with associated PRRs to signal defence responses (Hander et al., 2019; Wang, W. et al., 2020). Ca^{2+} signals are decoded by several groups of Ca^{2+} binding proteins which fall into 4 distinct groups: Calmodulin (CaM), Calmodulin like proteins (CaM-like), Calcium-dependent protein kinases (CPKs) and calcineurin b-like proteins (Ranty, Aldon and Galaud, 2006; Kudla, Batistič and Hashimoto, 2010; Poovaiah et al., 2013; Yuan et al., 2017; La Verde, Dominici and Astegno, 2018; Yip Delormel and Boudsocq, 2019). Calmodulin proteins have already been implicated in decoding MTI signalling due to their binding affinity with CNGC Ca^{2+} channels e.g. CaM7 is believed to block the CNGC2 and CNGC4 Ca^{2+} channel and becomes active following phosphorylation of the channel by BIK1, in addition CaM2 has been shown to interact with CNGC19 (Fischer et al., 2013; Meena et al., 2019; Tian, W. et al., 2019). Ca^{2+}signals are not only implicated in MTI but also in ETI responses, in particular CPKs have been identified as Ca^{2+} decoding proteins during ETI responses

and have been shown to regulate WRKY transcription factors during pathogen invasion (Gao, X. et al., 2013; Liu et al., 2017; Yip Delormel and Boudsocq, 2019). On top of this, Ca^{2+} responsive transcription factors such as Calmodulin binding transcription activator (CAMTA3) have been shown to regulate key genes that are involved in ETI pathways such as *EDS1* and pathogen related early response genes are enriched for CAMTA transcription binding motifs (Jacob et al., 2017; Kim et al., 2017; Lolle, Stevens and Coaker, 2020). Therefore, Ca^{2+} signalling is increasingly being found to co-ordinate both MTI and ETI responses and may provide the linchpin that connects MTI and ETI to their downstream signalling components.

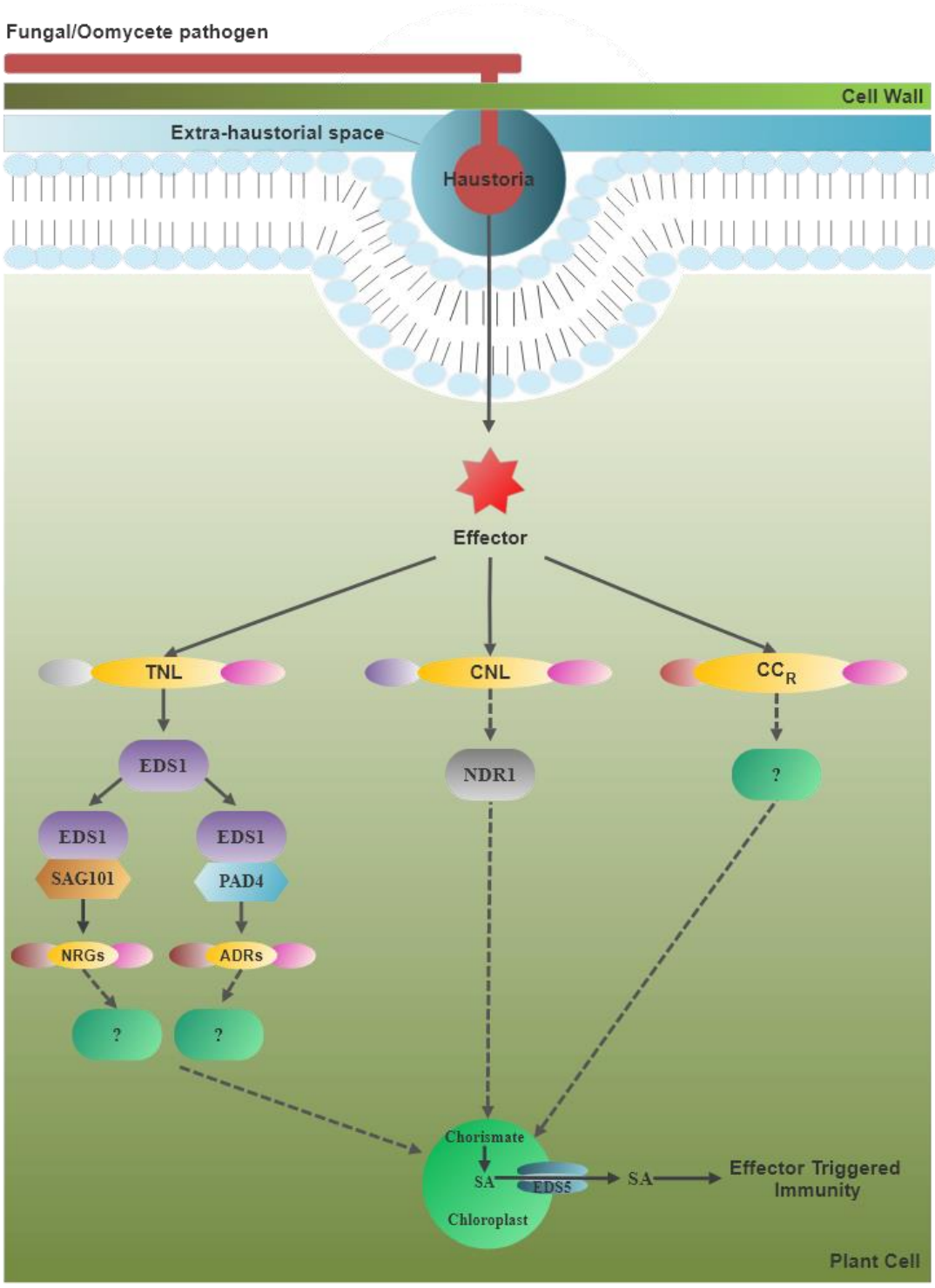


Figure 1.4: NLR signalling, the known components
Model showing the downstream signalling components utilised by TNLs, CNLs and CC_R NLRs following infection by fungal or Oomycete pathogens that lead to the production of Salicylic acid (SA) in the chloroplast. Figure adapted from (Lapin et al., 2019).

Structural insights into NLR activation

One area of research that frustrated plant molecular immunologists for years was the study of how NLRs are mechanistically activated. Many attempts have been made to structurally analyse plant NLRs but due to the hydrophobicity of the LRR domain, obtaining crystals of these proteins was particularly challenging. The first crystal structures published were partial structures, solely of the TIR domains of two TNLs Resistance to powdery mildew 4 (RPS4) and Resistance to *Ralstonia solanacearum* (RRS1) that operate as a paired NLR system (Williams, S.J., 2014). These crystal structures showed the formation of a TIR-TIR heterodimer at an interaction interface between two alpha helices (αa and αe), which if mutated abolished the HR triggered during the application of the avirulence protein, a finding that has been confirmed by several later studies (Williams, S.J., 2014; Williams, S. et al., 2016; Zhang, Y. et al., 2016; Newman et al., 2018). The formation of TIR-TIR homo or heterodimers was therefore determined to be crucial for TNL signalling. More recently, Cryo-EM structures of CNL HOPZ activated disease resistance 1 (ZAR1) were obtained, providing much needed evidence pertaining to the mechanistic structural reorientation of NLRs following infection by phytopathogens, shown in Fig 1.5 (Wang, J., Hu, et al., 2019; Wang, J., Wang, et al., 2019). ZAR1 is a resistance gene able to recognise bacterial pathogen *Xanthomonas campestris* via the decoy model of NLR activation (Wang, G. et al., 2015). *X. campestris* secrets an effector AvrAC into *A. thaliana* cells to uridylate a key MTI signalling component BIK1 that is induced following the binding of flg22 to the PRR and co-receptor complex FLS2-BAK1 (Lu et al., 2010; Wang, G. et al., 2015). *A. thaliana* contains a paralog of BIK1, PBS1-like protein 2 (PBL2) which acts as a decoy protein and is also uridylated by *X. campestris* effector AvrAC (Wang, G. et al., 2015). Uridylated PBL2 binds to ZAR1s RLCK pseudokinase partner RKS1 which forms a complex with the ZAR1 LRR repeat domain. The interaction of uridylated PBL2 with RKS1 subsequently causes a conformational change of ZAR1, releasing the autoinhibition of the NLR by its LRR domain (Wang, J., Wang, et al., 2019). The conformational change induced by this interaction then ejects ADP from the NB-ARC domain, replacing it with ATP to activate the NLR (Wang, J., Hu, et al., 2019; Wang, J., Wang, et al., 2019). Furthermore, active ZAR1 then proceeds to pentamerize through the oligomerization of the ZAR1 CC domain forming a ‘resistosome’ structure that has structural

similarities to the apaf1 and CED-4 apoptosomes and the inflammasome complex formed by animal NLRs such as NLRC4 (Wang, J., Hu, et al., 2019). Both apoptosomes and inflammasomes recruit and activate caspases upon formation resulting in apoptotic cell death (Li, Yini et al., 2017). Plants have no current known homologs of caspases, therefore the mechanism by which the resistosome functionally activates cell death is likely to be different to the apoptosomes and inflammasomes, although there are some caspase inhibitors that can prevent cell death responses in plants (Kabbage et al., 2017). Interestingly, the Cryo-EM structure of the active ZAR1 pentamer revealed the release of a 'funnel-shaped structure' formed of the initial α-helices from each of the five ZAR1 proteins, that could form an ion channel if inserted into a membrane (Wang, J., Hu, et al., 2019). If the formation of an ion channel could be proved, then this would provide a potential signalling mechanism which resistosome structures could employ to signal cell death.

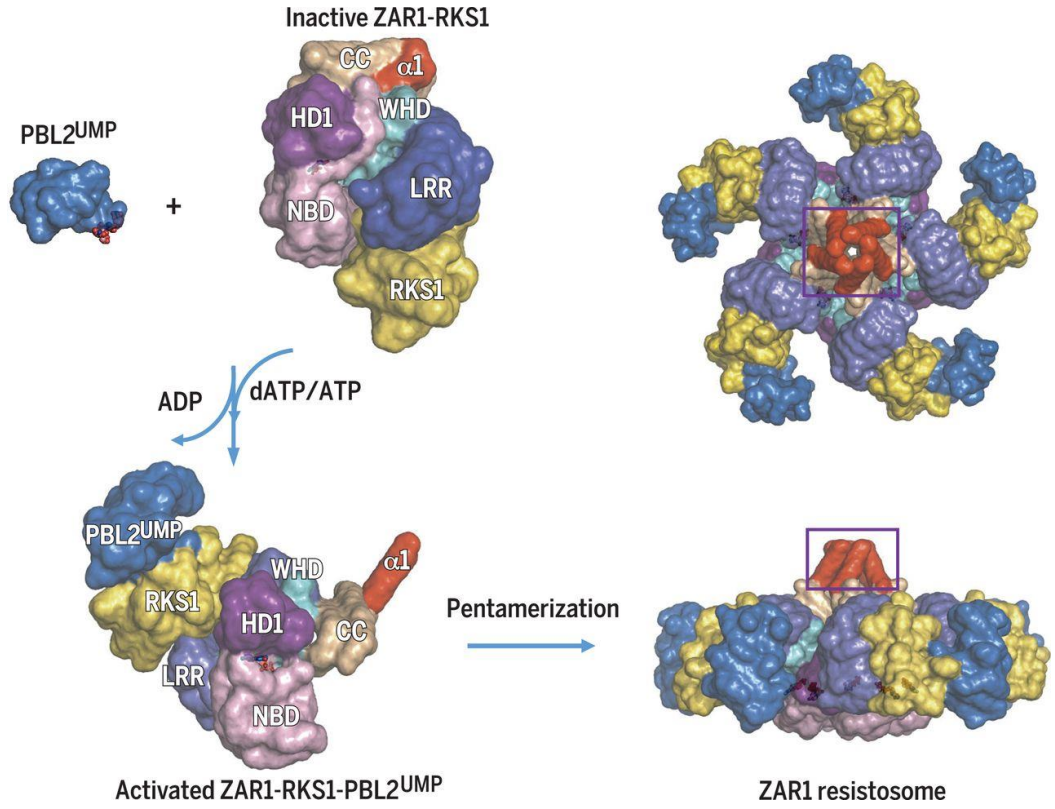


Figure 1.5: ZAR1 structure and activation
Model of ZAR1 activation reproduced from (Wang, J., Hu, et al., 2019). Following infection by bacterial pathogen *Xanthomonas campestris pv campestris* the plant decoy protein PBL2 becomes uridylated by *X. campestris* effector AvrAC. Uridylated PBL2, then binds to RKS1 which is in association with ZAR1, casuing the exchange of ADP for ATP, activating ZAR1 which then pentamerises to form the resistome complex.

Pairing up – the emergence of dual NLR systems and the sensor-helper model

Effector-NLR interactions have been heavily investigated since their discovery as a plant-pathogen interaction interface. Originally, the interaction of effector and NLR was believed to operate by a gene for gene hypothesis, whereby each NLR would have a single corresponding effector (Flor, 1971). This hypothesis then evolved as NLRs where identified that could recognise multiple pathogen effectors (Jones, J., D. G. and Dangl, 2006). More recently, our knowledge of the diversity of NLR activation has expanded with the identification of NLRs that require partner NLRs to function (Bialas et al., 2017). The most heavily studied paired NLR systems are RPS4 and RRS1 from *A. thaliana* and RGA4 and RGA5 from *Oryza sativa* (Bailey et al., 2018). These NLRs are often located in a tandem inverse arrangement in the genome, sharing a promoter region and are therefore believed to be co-regulated (Okuyama et al., 2011; Césari et al., 2014; Le Roux et al., 2015; Bailey et al., 2018). Tandemly orientated NLRs operate via the sensor-helper model whereby one NLR that usually contains an integrated domain operates as the sensor and the other executes or 'helps' activate defence signalling. The integrated domain containing NLR is referred to as the sensor as the integrated domain has been shown to interact with pathogen effector proteins, effectively acting as a bait allowing the plant to 'sense' the presence of the pathogen through the presence of the pathogens effectors (Narusaka et al., 2009; Sarris, Panagiotis F. et al., 2015). The type example of this is the RPS4 and RRS1 system, where both NLRs form a heterodimer in the cell prior to infection, and the interaction of the sensor NLR with the effector causes immune signalling by the executioner NLR (Huh et al., 2017). In addition to these two systems, several other paired NLR systems have been discovered, that share this spatial genetic arrangement, including TNLs, CSA1 and CHS3 and CNLs, Pik-1 and Pik-2 (Ashikawa et al., 2008; Xu et al., 2015). The presence of tandemly arranged paired NLR systems in both TNLs and CNLs shows that this is a conserved mechanism that some NLR systems utilise to bring about defence responses. However, not all paired NLR systems are found in this tandem arrangement in the genome. Some 'paired' systems form a network of sensor NLRs that signal through helper NLR nodes, such as the NLRs required for cell death (NRC) family of NLRs in the Asterids where NRC4 is required for ETI signalling conferred by multiple

helper NLRs against Oomycete, nematode and insect pathogens (Wu, C.-H. et al., 2017). Sensor NLRs commonly contain extra integrated domains that facilitate the recognition of pathogen effectors, revealing a common evolutionary trend that allows hosts to recognise a pathogens presence through an integrated domain (van Wersch and Li, 2019). Therefore, integrated domains in NLRs offer the enticing prospect for potential engineering approaches, whereby novel recognition capabilities could be designed into existing NLR architectures.

Integrated domains: How to snare an effector

Integrated domain containing NLRs are being identified across the plant kingdom and are becoming increasingly interesting due to their potential to inform NLR engineering approaches to developing disease resistance (Jones, J.D.G., Vance and Dangl, 2016; Kroj et al., 2016; Sarris, P. F. et al., 2016; Bailey et al., 2018; Stein et al., 2018). In order to be of use for resistance gene engineering, integrated domains need to be able to 'sense' a pathogens presence either through direct or indirect association of an effector with the integrated domain. If this association of an integrated domain with a pathogen molecule can be detected then the integrated domain can be considered as a suitable sensor domain that could be incorporated into an NLRs core architecture. The potential of each integrated domain depends on its scope of recognition, as some NLR sensor domains can recognise multiple pathogen effectors (Cesari et al., 2013).

The amount of NLRs containing IDs varies between 1-15% of NLRs in a single species but they are present in almost all plant genomes that have been analysed (Kroj et al., 2016; Sarris, P. F. et al., 2016). Some NLR IDs occur in greater abundance than others, for example the WRKY domain found in a family of transcription factors that are responsive to pathogen invasion has been identified in the architectures of NLRs in 13 different species and in at least 35 different NLRs, whereas others are still being identified in new studies and may only be present in one NLR (Dong, J., Chen and Chen, 2003; Sarris, P. F. et al., 2016; Van de Weyer et al., 2019). In addition to WRKY domain containing NLRs, the *Oryza sativa* R-gene RGA5 has been shown to contain a RATX1 NLR-ID which is similar to a copper chaperone protein structure, the RATX1 domain acts as a bait for *Magnaporthe oryzae* effectors AVR1-CO39 and AVR-Pia causing defence activation with its NLR partner RGA4 (Cesari et al., 2013; Césari et al., 2014; Ortiz et al., 2017). Therefore,

IDs have been shown to be the putative targets of pathogen effectors such as the WRKY domain or the RATX1 domain. The abundance of particular IDs in multiple NLRs suggests common targets of pathogen effectors and provides an important insight into potential domains that could be utilized for engineering future resistance genes.

The LIM domain as an integrated domain

One domain that has been identified that occurs in relative abundance as an NLR ID is the Lin-11, Isl-1 and Mec-3 (LIM) domain (Sarris, P. F. et al., 2016). The LIM domain is present in all eukaryotes, it is comprised of two zinc finger motifs separated by a two amino acid spacer and is involved in protein-protein interactions (Schmeichel and Beckerle, 1994; Zhao, M. et al., 2014). In plants there are four categorised protein families containing LIM domains, three of the four families contain two LIM domains directly adjacent to each other, the fourth group of LIM domain proteins are the plant specific DA1 and DA1 related (DAR) proteins which are comprised of only one LIM domain and are attached to previous domain of unknown function DUF3633, which has recently been shown to have metallopeptidase activity (Eliasson et al., 2000; Dong, H. et al., 2017). It is the DA1 family of LIM-Peptidase domain containing proteins that is associated with NLRs as the integrated domain, and has been identified in *A. thaliana* NLR Chilling sensitive 3 (CHS3) and NLR-like protein DAR5 as well as TNLs and CC_Rs from *Malus domestica*, *Prunus persica* and *Medicago truncatula* (Sarris, P. F. et al., 2016). In this thesis, I show that two resistance genes, both encoding LIM-Peptidase integrated domains cause resistance to *Albugo candida* in *Arabidopsis thaliana*. Therefore, the LIM-Peptidase domain is a likely target of pathogen effectors.

Current research surrounding the DA1 family has focused on its role in seed and organ development, in *A. thaliana* and *Brassica napus*, where *da1-1* dominant-negative mutants have been shown to increase leaf and seed size (Li, Yunhai et al., 2008; Wang, J.-L. et al., 2017). DA1 interacts and cleaves E3 ubiquitin ligases DA2 and Enhancer of DA1/Big Brother (EOD1/BB) as well as transcription factors TCP14 and TCP15 which are involved in controlling cell cycle associated proteins and is particularly associated with proteins controlling the endocycle (Xia, T. et al., 2013; Peng, Yuancheng et al., 2015; Li, N. and Li, 2016; Dong, H. et al., 2017). Microbes that interact with plants are known to induce localised endoreduplication,

these include Arbuscular mycorrhizal fungi, biotrophic fungal pathogens such as powdery mildew, nematodes and even viruses (Wildermuth, M.C. et al., 2017). This insight into plant-microbe interactions as well as the identification of two resistance genes containing LIM-Peptidase domains that are active against *A. candida* suggests that *A. candida* is interacting with endoreduplication associated proteins such as DA1 to increase its nutritional uptake from the plant upon infection and this interaction is exploited by the plants immune system to detect the presence of the pathogen.

Albugo candida and white rust resistance

Several important phytopathogenic organisms are found in the Stamenopila kingdom, particularly from the Oomycete class, including *Phytophthora infestans* (Potato late blight disease), *Hyaloperonospora arabidopsis* (downy mildews) and *Albugo candida* (White Rust) (Wang, Yan, Tyler and Wang, 2019). *A. candida* is a Brassicaceae infecting Oomycete phytopathogen that has an obligate biotrophic lifestyle. It has a wide host range, infecting over 200 Brassicaceae species including vegetable and oilseed crops as well as the model plant species *A. thaliana* (Saharan et al., 2014). *A. candida* reproduces both sexually and asexually, releasing both asexual zoospores and sexual Oospores. During the asexual life cycle of *A. candida*, zoosporangia form in pustules adhered to the abaxial surface of host leaves, these zoosporangia then dehisce in water releasing flagellated zoospores which swim chemotactically to host stomata where they enter the host (Holub et al., 1995). Once inside the plant, hyphae develop and penetrate into the mesophyll layer where haustoria are formed that penetrate through the host cell wall forming a nutrient exchange layer between the host cell wall and plasma membrane, called the extra-haustorial space, into which effectors are secreted and nutrients are uptaken. Finally, after ~7 days *A. candida* forms pustules on the plant tissues and the asexual life cycle starts again.

One unusual feature of *A. candida* infection is its ability to cause strong immunosuppression, resulting in an immunocompromised host. Immunocompromised hosts are susceptible to secondary infections and can subsequently be colonised by non-host pathogens for example *Phytophthora infestans and Hyaloperonospora arabidopsis* can infect *A. thaliana* or *Brassica juncea* respectively following pre-inoculation by *A. candida* (Cooper et al., 2008;

Prince et al., 2017). Immunosuppression caused by *A. candida* is particularly problematic on crop plants because crops not only lose yield due to growth defects caused by the *A. candida* infection but can subsequently succumb to secondary infection resulting in yield losses of up to 90% (Saharan and Verma, 1992; Saharan et al., 2014). Although *A. candida* is identified as a single species, it contains several independent physiological races that specialise on separate hosts species (Borhan, M. Hossein et al., 2008; McMullan et al., 2015; Jouet et al., 2019). In addition to this, some races show differential intraspecific infection phenotypes, for example *A. candida* race 4 isolate AcEM2 is able to infect *A. thaliana* accession Ws-3 but not Col-0 (Borhan, M. Hossein et al., 2008). The ability of *A. candida* to differentially infect distinct host ecotypes, has recently been determined to be due to the intraspecific variation in NLR distribution between *A. thaliana* populations resulting in a pathosystem that is constantly in flux (Cevik et al., 2019). The difference in phenotypic response to specific isolates of *A. candida* means that we can determine the causal genes underpinning resistance by generating recombinant inbred lines (RILs) and utilise phenotype-based mapping studies to pinpoint the genetic loci associated with resistant responses.

NLRs that are active against *A. candida* are collectively known as *White Rust Resistance (WRR)* genes. There are currently 5 published *WRR* genes found in *A. thaliana* accessions: *WRR4A* and *WRR4B* (Col-0), *WRR8* (Sf-2), *WRR9* (Hi-0) and *WRR12* (Ler-0) that confer resistance to different *A. candida* races (Borhan, M. Hossein et al., 2008; Cevik et al., 2019). The presence of differing numbers of *WRR* genes between *A. thaliana* populations highlights that an arsenal of NLRs are required within populations to provide a species wide pan-NLRome conferring resistance against a single phytopathogen (Cevik et al., 2019). Therefore, NLRs have to be rapidly evolving to generate a large enough repertoire to combat all the disease threats posed by a multi-pathogen ecosystem. Increasing our understanding of the quantity of NLRs required in one species to be active against one pathogen will help inform breeding approaches in the future, particularly into the number of resistance genes needed to be employed in crop populations to provide durable long-term resistance against a single pathogen.

Thesis Aims

The research aims of this thesis are based on previous work of Cevik *et al,* (*in prep*) that multiple resistance phenotypes were identified in *A. thaliana* Recombinant inbred lines (RILs) derived from a cross of resistant *A. thaliana* ecotype Col-5 with susceptible ecotype Ws-2, when challenged with *A. candida* race 4 isolate AcEM2. A Quantitative trait loci (QTL) analysis performed on the RILs revealed three loci associated with AcEM2 resistance, one of which corresponds to *WRR4* a previously identified *WRR* gene (Borhan, M. Hossein et al., 2008) and two additional loci termed *WRR5* and *WRR7* that were associated with the resistance phenotypes. In this thesis I aim to, identify and characterise the response of resistance genes *WRR5A*, *WRR5B* and *WRR7* to AcEM2 and determine whether they are the causal agents of resistance to AcEM2 in *A. thaliana*. Both *WRR5B* and *WRR7* encode integrated LIM-Peptidase domains as part of their protein architecture. Therefore, I will then determine the role this domain plays in *A. thaliana* immunity/susceptibility to *A. candida*. I will also attempt to engineer novel resistance genes in *A. thaliana* against *A. candida* using domain swapping experiments of the WRR7 integrated LIM-Peptidase domain with other highly similar LIM- peptidase domains obtained from other *A. thaliana* proteins to determine how specific integrated domains need to be in order to generate resistance. This information will inform future attempts to breed novel resistance genes using the integrated decoy model of NLR derived resistance.

Chapter 2

Materials and Methods

Biological Material

A. thaliana lines were sown on F2+ S compost (Scotts, UK) or ½ Murashige and Skoog basal medium (MS) (Sigma-Aldrich, USA) and cold stratified at 4°C for 2-4 days. If transgenic seeds were under Kanamycin selection, they were sown on MS plates containing 50 µg/ml of Kanamycin. Seedlings were grown on MS plates grown for ~2 weeks then transplanted into F2+ S compost (Scotts, UK) and kept under propagators for 48 hours. Seeds sown on soil were germinated in stock pots under propagators for 2 weeks. Once seedlings showed their first set of true leaves, seedlings from the stock pots were transplanted into fresh F2+ S compost and propagators were removed 48 hours after transplanting. All seedlings were germinated in a controlled environment room maintained at 21°C under a 10-hour day (light intensity 100 µmol/m^2) and 14-hour dark cycle. If seeds were required from the plants, then they were moved to long day conditions (16-hour day and 8-hour dark cycle) after ~4 weeks growth.

Nicotiana benthamiana and *Nicotiana tabacum* seeds were sown on M2 soil (Scotts, UK) and grown under a propagator in a controlled environment incubator maintained at 22°C under a 14-hour day and 10-hour dark cycle. Seedlings were transplanted into fresh M2 soil 2 weeks post sowing and kept under a propagator for 1 week after transplanting.

A. candida races were maintained on 4-week-old susceptible *A. thaliana* lines (either Ws-2 or Col-*eds1.2-wrr7*). *A. candida* was propagated by collection of infected *A. thaliana* leaves, suspension of zoosporangia in ice chilled water and filtration of the suspension through a single layer of Miracloth ™ (Merck Milipore, Germany). The zoospore containing filtrate was kept on ice and then sprayed onto fresh susceptible plants using a pressurised spray gun. After inoculation, plants were kept in a 4°C cold room for ~12 hours under a propagator before transfer to a controlled environment cabinet with a day regime of 10 hours light at 22°C and a night regime

of 14 hours dark at 18ºC, propagators were removed after ~1 day. This process was repeated every 7-10 days once pustules had developed.

E. coli strains DH5-α or DH10B transformed with a plasmid containing different antibiotic resistance genes was grown on ½ salt Liquid broth (LB) (5% NaCl, 10% Bacto-peptone, 5% yeast extract) plates containing 1.5% agar and appropriate antibiotics, plates were incubated in the dark at 37ºC for 12 hours and stored at 4ºC for further use. Cultured *E. coli* was grown in 5ml ½ salt LB containing (50µg/ml of Kanamycin, 100 µg/ml Carbenecillin, 100 µg/ml Spectinomycin) at 37ºC for 12 hours and shaken at 180rpm.

A. tumefaciens strains were maintained in 20% glycerol stocks at -80ºC. Strains were re-inoculated on ½ salt LB plates containing 1.5% agar and appropriate antibiotics (typically 50 µg/ml of Rifampicin and Kanamycin and 20 µl/ml Gentamycin) and grown at 28ºC for 12-48 hours.

Trypan Blue staining

The trypan blue staining protocol was adapted from Fernández-Bautista et al (2016). Leaf samples were harvested immediately before staining and placed in glass universals and submerged in Trypan Blue staining solution: 50% lacto-phenol Trypan Blue (25% phenol, 25% glycerol 25% lactic acid (85% W/W), 25% water and 10mg/ml trypan blue dye) with 50% ethanol. The universals containing the sample and staining solution were then heated in a glass beaker containing water and the samples were incubated for 1 minute. The staining solution was then discarded and replaced with Chloralhydrate (14 M) and incubated at room temperature for 24-48 hours, the Chloralhydrate was then removed and replaced with fresh Chloralhydrate and left for a further 12 hours and then replaced with 50% glycerol prior to imaging.

Cloning of *WRR5A* and *WRR5B*

To clone the genomic fragment containing both *WRR5A* (*At5g178880*) and *WRR5B* (*At5g17890*), TAC clone JAtY79I19 obtained from *Arabidopsis* accession Col-0 was first digested with *Kpn*I-HF (NEB) generating a 20,272 bp fragment containing both genes that was purified from a 1 % agarose gel. The purified DNA fragment was then ligated into *Kpn*I-HF (NEB) digested and dephosphorylated *pCambia2300* vector and electroporated in to DH10B cells. As we used single restriction enzyme,

our cloning yielded constructs with *WRR5A* and *WRR5B* in two different orientation (i.e. pCambia2300:*TermWRR5A:WRR5A:proWRR5A&B:WRR5B:WRR5BTerm* or *TermWRR5B:WRR5B:proWRR5B&A:WRR5A:WRR5ATerm*). These two distinct constructs were then digested with *Sal*I (NEB) and digested plasmids with the inserts were run on 1% agarose gel. Plasmid DNA with 7,959 bp fragment harbouring *WRR5A* only or 12,313 bp fragment containing *WRR5B* only was then isolated from the gel, purified, self-ligated and transformed into DH10B cells. Resulting plasmids containing only *WRR5A* or *WRR5B* genomic fragments were then confirmed by sequencing.

USER Cloning

Genomic or cDNA fragments of target genes or gene fragments were PCR amplified from DNA using the KAPA long range hot start PCR kit (KAPA biosystems, Germany), genomic fragments were amplified from *A. thaliana* (Col-0 ecotype) genomic DNA or *A. thaliana* DNA in JAtY clone libraries, cDNA fragments were amplified from cDNA synthesised from RNA extracted from *A. thaliana* Col-0 plants. A list of gene targets and primers used can be found in the Table 2.1. Fragments were then PCR purified or Gel extracted depending on the presence of non-specific bands or primer dimers, using QIAquick PCR Purification Kit or QIAquick Gel Extraction Kit (Qiagen, Germany). Fragments were then fused in a Uracil-specific excision reagent (USER) reaction (Nour-Eldin, Geu-Flores and Halkier, 2010). Whereby DNA fragments were fused in equimolar concentrations by incubating in 10 µl USER reactions at 37^{o}C for 20 minutes followed by 25 minutes at 25^{o}C, reactions contained 1 µl 10x Cut Smart buffer (New England biolabs, USA), 1 µl USER enzyme, 1 µl of the desired vector at a concentration of 30 ng/µl and DNA fragments that were diluted so that the vector to amplicon ratio was 1:5 and then made up to 10 µl with sterile MQ water. Amplicons were then fused by adding 1 µl T4 DNA ligase (New England Biolabs, USA) and 1.2 µl T4 DNA ligase buffer (New England Biolabs, USA) to the USER reactions and incubating at 16^{o}C for a minimum of 1 hour. Cloned genes driven by their native promoters were cloned into the pUSER LBJJ233 vector whereas genes targeted for overexpression were cloned into the pUSER LBJJ234 containing the *35S* cauliflower mosaic virus promoter and *Octopine synthase* terminator (Fig 2.1). Genes that were C-terminally tagged had

their stop codons removed and mixed with tag DNA fragments containing a reading frame encoding Glycine-Serine spacer (GSGS) followed by the intended tag (see table for tag and primer list). Ligated vectors containing the target gene constructs were subsequently chemically transformed or electroporated into chemical or electro competent DH10B or DH5-α *E.coli* cells (New England Biolabs, USA) and grown on ½ salt LB 1.5% agar plates containing 50 µg/ml kanamycin for ~12 hours at 37^{o}C. Colonies were screened using PCRs to select the colonies with the desired insert, cultured overnight in ½ salt LB containing 50 µg/ml kanamycin at 37^{o}C and shaken at 180 rpm. Plasmids were extracted using QIAprep Spin Miniprep Kit (Qiagen, Germany) and sequenced using Sanger sequencing (Eurofins, Luxembourg). Correct constructs were then transformed into *Agrobacterium tumefaciens* strain GV3101 (containing pMP90) and grown on ½ salt LB 1.5% agar plates containing 50 µg/ml of Rifampicin and Kanamycin and 20 µl/ml Gentamycin at 28^{o}C for 48 hours. Colonies were cultured for 12 hours in ½ salt LB containing 50 µg/ml Rifampicin and Kanamycin and 20µg/ml Gentamycin at 28^{o}C shaking at 180 rpm before being stored in 20% glycerol stocks stored at -80^{o}C.

Floral dipping

A. thaliana lines were transformed using the floral-dip method (Bent, 2006). Glycerol stocks of *A. tumefaciens* strain GV3101 containing the construct of interest were streaked onto ½ salt LB media plates containing 1.5% agar and appropriate antibiotics and grown for 24 hours at 28^{o}C. Colonies were then suspended in 250 ml ½ salt liquid LB cultures containing the same antibiotic cocktail as used on the LB plates and grown for 12 hours at 28^{o}C at 180 rpm. Liquid cultures were centrifuged at 5000 xg at room temperature for 15 minutes, the supernatant was discarded, and the pellet was resuspended in 5% sucrose solution containing 0.05% Silwet L-77. *A. thaliana* secondary bolts were dipped for ~15 seconds in the sucrose solution containing the *A. tumefaciens* GV3101 strain containing the gene construct of interest in the Ti plasmid. Dipped plants were then kept in the dark for 12 hours then removed and grown at 21^{o}C under a 16-hour day and 8-hour dark cycle. Transgenic seeds were subsequently selected using the fast red selection method (Shimada, Shimada and Hara-Nishimura, 2010).

DNA extraction and Genotyping

DNA for genotyping or PCR was extracted using the CTAB method (Porebski, Bailey and Baum, 1997). Leaf samples were ground in 100 µl of DNA extraction buffer (0.14M d-Sorbitol, 0.22M Tris-HCl pH 8, 0.022M EDTA pH 8, 0.8M NaCl, 0.8% CTAB (Cetrimonium bromide), 0.1% n-Laurylsarcosine) containing 1µl RNase A and a little bit of sterile sand per sample. Samples were then vortexed for 15 seconds and incubated at 65^{o}C for 5 minutes. Following incubation 100 µl of Chloroform was added to each sample and they were centrifuged at 13,000 xg for 5 minutes. The upper phase of the supernatant was transferred to a new tube and DNA was precipitated by adding 100 µl of iso-propanol and incubated for 15 minutes at room temperature. DNA was then pelleted by centrifugation at 13,000 xg for 15 minutes, the supernatant was removed, and the pellet was washed twice with 1 ml of 70% ethanol and resuspended in Tris-HCl pH 8 and stored at -20^{o}C.

For genotyping DNA <1 kb in size DNA was extracted using Chelex (Biorad): two leaves were collected from target plant and suspended in 250 µL of 10% Chelex, samples were kept on ice. The leaf was then crushed using a sterile pipette tip before vortexing twice for 5 seconds. Samples were then incubated at 96^{o}C for 5 minutes, vortexed a second time before being incubated a further 5 minutes at 96^{o}C. Samples were then vortexed a further 3 times before being left on ice for 10 minutes, vortexed for a final time before spinning down briefly and the supernatant was taken for use in PCR.

DNA for whole genome sequencing or re-sequencing was extracted using the DNeasy plant MaxiPrep Kit (Qiagen, Germany) and Plasmids were extracted using QIAprep Spin Miniprep Kit (Qiagen, Germany).

PCR reactions using DNA extracted by the Chelex method were prepared in a 20 µl reaction containing 2.5 µl DNA sample, 2 µl 10x Dream Taq buffer, 0.2 µl 10 mM dNTP mix, 0.1 µl 10x Dream Taq DNA polymerase, 1 µl of each 10 mM primer (Table 2.1) and 13.2 µl MQ water. PCR reactions performed using DNA extracted using the CTAB method was performed in 10 µl total reactions, containing 1 µl DNA sample, 1 µl 10x Dream Taq buffer, 0.2 µl 10mM dNTP mix, 0.1 µl 10x Dream Taq DNA polymerase, 1 µl of each 10 mM primer and 5.5 µl MQ water. Amplifications of DNA fragments using USER primers for USER cloning was

performed in 50 µl total reactions using 1 µl DNA sample, 25 µl KAPA HiFi U^+ master mix (KAPA biosystems, Germany), 2.5 µl of each 10 mM primer and 19 µl MQ water. PCR was performed using a SimpliAmp thermal cycler (Thermo-fisher), the PCR cycle was programmed for initial denaturation at 94ºC for 3 minutes followed by 35 cycles of denaturing at 94ºC for 30 seconds proceeded by annealing at 56ºC for 30 seconds and then extension at 72ºC for 2 minutes. Once the PCR cycles were complete, samples were cooled at 15ºC for 15 seconds then maintained at 20ºC until use. 2 µL of each sample were then loaded into 2% agarose gel in 1x TAE buffer containing 2.5µL of ethidium bromide, gels were run at 80 v for~40 minutes and then imaged using a UVP imager (Analytic Jena, Germany) and VisionWorks image acquisition and analysis software (Analytik Jena).

RT-qPCR

RNA was extracted using Direct-zol RNA miniprep plus kit (Zymo research) and frozen at -80ºC. cDNA was synthesized from RNA using SuperScript™ IV First-Strand Synthesis System (Invitrogen) using 1-2 µg total RNA with oligo dT primers and frozen at -80ºC prior to use.

cDNA samples were diluted in sterile milliQ water at 1 in 5 times dilution. Each 20 µl reaction contained 1.5 µl diluted cDNA, 10 µl 2x MyTaq HS ready mix (Bioline), 1 µl 20x Eva green dye (Biotium), 0.4 µl each primer (10mM) and 6.7 µl MQ water. PCR reactions were run on Aria Mx PCR system (Agilent) under the following conditions: initial denaturation 95ºC for 2 minutes followed by denaturation for 5 seconds, annealing 57ºC for 30 seconds repeated for 45 cycles. A melt cycle was performed on each primer set: denaturation at 95ºC for 30 seconds followed by annealing at 65ºC for 30 seconds followed by a further denaturation step at 95ºC for 30 seconds to check primers were amplifying a single fragment.

Data was analysed using the ΔΔCT method, expression data of the gene of interest was normalised against housekeeping gene Protein phosphatase 2a subunit A3 (PP2AA3) (Hong et al., 2010).

Transient expression

Agrobacterium tumefaciens strain GV3101 glycerol stocks containing target genes were streaked onto on ½ salt LB plates containing 1.5% agar and appropriate antibiotics at 28ºC for 12 hours. The bacteria were then grown in 10 ml cultures of ½

salt LB containing appropriate antibiotics for 12 hours at 28°C and shaking at 180 rpm. Bacteria cultures were then centrifuged at 3000 rpm (20°C) for ~7 minutes and resuspended in 5-10 ml infiltration buffer (10 mM $MgCl_2$ + 10 mM MES buffer adjusted to pH 5.6) depending on the size of the pellet the OD_{600} of each culture was then measured from a 1 in 10 dilution of the bacterial culture. The OD_{600} of each bacterial culture was then adjusted to the desired OD_{600} in the final inoculum, typically an OD_{600} of 0.5 was used unless the construct was known to have a particularly high expression after infiltration. To aid in the expression of the agro-infiltrated constructs the Tomato bushy stunt virus protein P19 was co-infiltrated into *N. benthamiana* leaves at an OD_{600} of 0.2 (Canto et al., 2006). 150 µM Acetosyringone (Sigma-Aldrich, USA) was then added to each inoculum and they were left to incubate at room temperature for 1 hour. Transient expression was carried out in 4-week-old *Nicotiana benthamiana* or *Nicotiana tabacum* leaves. Each inoculum was infiltrated into 2-3 leaves by pricking the abaxial surface of the leaf with a needle and injecting the inoculum into the leaf using a syringe. Infiltrated plants were then left for 2-3 days before imaging or being frozen in liquid nitrogen for protein analysis.

Protein sub-cellular localisation and microscopy

Fluorescently tagged target proteins were transiently expressed in *Nicotiana benthamiana* and leaf sections were visualised 2 days post infiltration. Leaf sections were infiltrated with water before mounting on microscope slides and imaged using a Nikon eclipse 90i confocal microscope, using two lasers, one at 488 nm and one at 543 nm generated by the Nikon D-eclipse C1 confocal microscope system and controlled using the EZ-C1 microscope software (Nikon). Any proteins thought to be plasma membrane localised were observed from plasmolysed leaf tissue. Leaves were plasmolysed by infiltrating 1 M sucrose solution into the leaf and leaving 5 minutes before being imaged.

Protein extraction

Leaf samples were collected in Liquid nitrogen and homogenised in a sterile pestle and mortar. The crushed sample was collected in a sterile tube and kept on dry ice until ready for use. Extraction buffer (150 mM Tris-HCl, (pH 7.5), 150 mM NaCl, 1 mM EDTA (pH 8), 1 mM EDTA, 10% Glycerol, 10 mM DTT, 0.2% Nonidet-40, 2% polyvinylpolypyrolidone and cOmplete™ EDTA free protease inhibitor tablets

(Roche diagnostics, Germany)) of an equal volume to the volume of homogenised tissue was added to the sample on ice. The sample and extraction buffer were then mixed by vortexing and incubated on a Belly shaker for 20 minutes at 4^{o}C. Cell lysates were then separated from the homogenised tissue by centrifugation at 5000 xg for 20 minutes at 4^{o}C, the supernatant was then filtered through a single layer of Miracloth ^{TM} (Merck Milipore, Germany) pre-saturated in MiliQ water. Extracted cell lysates were then mixed in a 2:1 ratio with 3x SDS loading dye (30% Glycerol, 3% SDS, 0.05% bromophenol-blue, 93.75 mM Tris-HCl (pH 6.8) and 20 mM DTT) and denatured by boiling at 95^{o}C for 10 minutes and stored at -20^{o}C.

Microsomal and Nuclear Fractionations

For the extraction of the microsomal fraction ~1 g of leaf tissue was ground in liquid nitrogen and homogenised in 2 ml pre-chilled sucrose buffer (0.33 M sucrose, 20 mM Tris-HCl pH 8, 1 mM EDTA, 5 mM DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF)). Homogenized sample was filtered through a 100 µm membrane and centrifuged at 2000 xg for 10 minutes at 4^{o}C. 200 µl of the supernatant was then mixed in a 2:1 ratio with 3x SDS loading dye and denatured by boiling at 95^{o}C for 10 minutes and stored at -20^{o}C and used as the total protein fraction. A further 200-300 µl of the supernatant was then centrifuged for 30 minutes at 100, 000 xg at 4^{o}C, the supernatant was then used as the soluble fraction and mixed in a 2:1 ratio with 3x SDS loading dye and denatured by boiling at 95^{o}C for 10 minutes and stored at -20^{o}C. The pellet was then suspended in 200 µl of sucrose buffer, to be used as the microsomal fraction and mixed in a 2:1 ratio with 3x SDS loading dye and denatured by boiling at 95^{o}C for 10 minutes and stored at -20^{o}C.

For nuclear protein extraction ~1 g of leaf tissue was ground in liquid nitrogen and lysed by adding 2 ml lysis buffer (20 mM Tris-HCl pH 7.5, 25% Glycerol, 20 mM KCl, 2 mM EDTA, 2.5 mM $MgCl_2$, 250 mM sucrose, 5 mM DTT, 1x cOmplete^{tm} protease cocktail inhibitor (Roche) and 1 mM PMSF) and homogenised by gentle agitation with a pipette. The homogenised tissue was then sequentially filtered through 100 µm and 40 µm nylon mesh. The total protein sample was then taken from filtrate and mixed in a 2:1 ratio with 3x SDS loading dye and denatured by boiling at 95^{o}C for 10 minutes and stored at -20^{o}C. Nuclei were then separated from

the remainder of the filtered homogenate by centrifugation at 1,500 xg at 4°C for 10 minutes. The Nuclei depleted fraction was then taken from the supernatant and mixed in a 2:1 ratio with 3x SDS loading dye and denatured by boiling at 95°C for 10 minutes and stored at -20°C. Pelleted nuclei were resuspended and washed in 3 ml Nuclei resuspension buffer (NRB: 20 mM Tris-HCl pH 7.5, 25% glycerol, 2.5 mM $MgCl_2$ and 5 mM DTT) containing 0.2% Triton X-100 and re-separated by centrifugation at 1,500 xg at 4°C for 10 minutes, the washing step was repeated a further 3 times. After the final wash the pelleted Nuclei were suspended in NRB without Triton X-100, spun down a final time suspended in 200 µl of NRB and mixed in a 2:1 ratio with 3x SDS loading dye and denatured by boiling at 95°C for 10 minutes and stored at -20°C.

Immunoprecipitation

For immunoprecipitation or co-immunoprecipitation 20 µL of antibody-bound beads, either ANTI-FLAG (Sigma-Aldrich, USA), Anti-V5 (Abcam, UK) or GFP-Trap (Chromotek, Germany) per sample were prepared by washing with 1 ml washing buffer (150 mM Tris-HCl, (pH 7.5), 150 mM NaCl, 1 mM EDTA (pH 8), 1 mM EDTA, 10% Glycerol, 10 mM DTT, 0.2% Nonidet-40, and cOmplete™ EDTA free protease inhibitor tablets (Roche diagnostics, Germany)). Washing buffer was removed following centrifugation at 7000 xg for 15 seconds in a centrifuge pre-cooled to 4°C.

Following washing, 1.5 ml of extracted cell lysates were incubated on ice with washed antibody-bound beads in LoBind™ tubes (Eppendorf, Germany) and incubated on a rotary mixer at 4°C for 2 hours. The beads were spun down at 7000 xg for 15 seconds in a centrifuge pre-cooled to 4°C and the cell lysate was removed, and the antibody beads were resuspended in with 1 ml washing buffer. Samples were then centrifuged and washed at least another 3 times. After the final wash residual washing buffer was removed using a syringe and needle (with 0.3 mm diameter) and the beads were mixed with 75 µl 3x SDS loading dye (30% Glycerol, 3% SDS, 0.05% bromophenol-blue, 93.75 mM Tris-HCl (pH 6.8) and 20 mM DTT) and denatured by boiling at 95°C for 10 minutes and stored at -20°C.

Immunoblotting

Extracted total protein (input) or Immunoprecipitated samples mixed with 3x SDS loading dye were denatured at 95^{o}C for 10 minutes prior to use. Protein samples were then run on 6, 8, 10 or 12% SDS-PAGE gels depending on protein size (generally proteins between 40-180 KDa were run on 8% gels, anything smaller were run on higher percentage gels) using a 10x Tris-glycine based buffer containing 1% SDS. Exceptionally large or small proteins (>200 KDa or <30 KDa) were separated on 4-16% RunBlue™ TEO-Tricine SDS Mini Gels (Expedeon, UK) in a 1x run Blue DDS Run Buffer (Expedeon, UK). Separated proteins were transferred from the acrylamide gels to nitrocellulose membranes (Bio-Rad, USA) in Trans-Blot Turbo ^{TM} 5x transfer buffer (Bio-rad), using a Trans-Blot Turbo ^{TM} protein transfer machine (Bio-Rad). Once the proteins had been transferred, membranes were washed once with 1x TBS (20 mM Tris-HCl (pH 7.5) and 150 mM NaCl) before being transferred to a blocking buffer (1x TBS, 0.1% Tween 20 and 5% non-fat milk powder) and incubated for 1 hour at room temperature on a belly shaker. Blocked membranes were then incubated in a blocking buffer containing primary antibodies Anti-FLAG-HRP (Sigma-Aldrich), Anti-V5-HRP (Abcam), Anti-GFP (Abcam), Anti-BAK1 (Agrisera, Sweden) or Anti-Histone H3 (Thermo Fisher, USA) at appropriate antibody concentrations (ranging from 1:5000 – 1:20,000 dilutions) for 12 hours at 4^{o}C on a belly shaker. If the primary antibody was not HRP conjugated then the membranes were transferred to another blocking buffer containing secondary antibody α-Goat anti-Rabbit IgG HRP (Thermo Fischer, USA) and incubated for 1 hr at room temperature. Following antibody treatment, the membranes were then washed three times with TBS-T (1x TBS + 1% tween 20) initially for 10 minutes followed by three further 5-minute washes. Membranes were then saturated with SuperSignal^{TM} west Pico plus (Thermo Fisher, USA), sometimes combined with SuperSignal^{TM} femto (Thermo Fisher) in a 10:1 ratio depending on the protein expression level, before developing on CL-Xposure^{TM} film (Thermo scientific). Membranes were stained in Ponseau S dye (0.1% w/v in 5% acetic acid) for 5 minutes on a belly shaker following development of the film.

IP-MS

IP-MS samples from lines overexpressing *WRR5A-V5* and *WRR5B-HF* as well as proteins from Ws-2 wild type plants were immunoprecipitated as above, however 30 µl of antibody beads were used to precipitate the proteins and immunoprecipitated samples were washed 5 times. Immunoprecipitated samples were then sent for liquid chromatography- mass spectrometry analysis at the Bristol proteomics facility. Proteins were identified using a Sequest search against the Uniprot *A. thaliana* Col-0 database and filtered using a 1% false discovery rate. Identified proteins from the sample database were then cross-referenced against the negative control database generated from the appropriate IP-bead control and proteins were only taken forward if they were present in the sample in greater than 5x the amount they appeared in the bead control.

EMS mutagenesis and Mutant analysis

Col-*eds1.2* seeds were chemically mutagenized using ethyl methanesulfonate (EMS) prior to the start of the project. Mutagenized seeds were grown, selfed and seeds were harvested from each individual M1 plant. ~500 seeds were then sown from each M2 (total of ~200,000 seedlings) line and screened with *Albugo candida* isolate AcEM2 after 2 weeks of growth. Susceptible seedlings were transplanted, treated with 25 mg/L metalaxyl, a fungicide known to also be active against Oomycetes (Sukul and Spiteller, 2000), selfed and harvested, generating M3 pools of seeds from each susceptible plant. ~100 seedlings from each pool of M3 seeds were rescreened with AcEM2 and M3 lines showing 100% susceptibility were selected for further analysis as the mutagenized gene of interest was homozygous. Selected M3 seed pools from each susceptible M2 line were then grown and backcrossed (BC) with Col-*eds1.2* to make the mutation of interest heterozygous in BC F1 plants. BC F1 plants were then grown to seed and BC F2 seeds were collected, sown, inoculated with AcEM2 after 3 weeks of growth and then screened. The segregation ratio of resistant: susceptible seedlings was analysed, with an expected 3:1 segregation ratio with a single causal gene. Susceptible individuals (~200 for each mutant line) were then selected and bulked for DNA extraction. DNA was extracted using the DNeasy plant Maxi prep kit (Qiagen, Germany) and bulked DNA was sent for whole genome

re-sequencing (Novogene, Singapore). Candidate causal mutations were then analysed using a direct sequencing approach (Sikora et al., 2012).

Bioinformatics

EMS mutagenesis candidate gene identification

Genome reads were obtained from bulked segregant populations from F_2 populations derived from a backcross of Col-*eds1.2* EMS mutant lines x Col-*eds1.2* as well as from Col-*eds1.2*. These reads were then analysed using the Simple mapping pipeline (SIMPLE) to identify candidate genes (Wachsman et al., 2017). The SIMPLE pipeline aligns sequence reads to the Col-0 reference genome, then identifies Single nucleotide polymorphisms (SNPs) between the mutant genome and the reference genome of interest (in this case Col-*eds1.2*) using the GATK haplotype caller. The effects of these SNPs were then identified using SnpEff software (Cingolani et al., 2012) and candidate genes are identified based on the SNP ratio and deemed to have a significant effect on a protein coding region (Wachsman et al., 2017). As well as the candidate list of genes, we analysed the output file containing all potential SNPs as in some cases a few contaminant reads can affect the segregant ratio and therefore don't show up in the candidate gene file.

Gene ontology analysis

Immunoprecipitation-mass spectrometry analysis generated potential interacting proteins with WRR5A and WRR5B. Gene ontology analysis was performed on the identified gene datasets using singular enrichment analysis in agriGO software v2.0 and compared to the *A. thaliana* TAIR 10 reference genome (Berardini et al., 2015; Tian, T. et al., 2017). Gene ontologies were determined to be significant at a P-value of <0.05%.

RNA sequence analysis

RNA was extracted from Col-*eds1.2-wrr7* seedlings infected with *A. candida* isolate AcEM2 or mock inoculated with water 2, 4- and 6-days post inoculation using Direct-zol RNA miniprep plus kit (Zymo research). RNA libraries were prepared by filtering mRNA from total RNA using poly-T oligos attached to magnetic bead substrates and then reverse transcribed into cDNA using SuperScript™ IV First-Strand Synthesis System (Invitrogen). cDNA was then sequenced by 150 bp paired end sequencing using the illumine HiSeq platform (Novogene). The transcriptome

for each sample was then constructed using a genome guided approach in TopHat v2.0.12, aligning the reads back to the *A. thaliana* Col-0 reference genome (Kim, D. et al., 2013). Read counts for each gene were subsequently generated in HTseq (Anders, Pyl and Huber, 2015). Both the TopHat and HTseq analysis was performed by Novogene and we received read count data from our cDNA libraries.

Read count data was then uploaded to iDEP 9.0 (Ge, Son and Yao, 2018) and analysed using a minimum counts per million (CPM) threshold of 0.5, the data was transformed using the edgeR log transformation with a pseudo count of 1 (Ge, Son and Yao, 2018). Differentially expressed genes (DEGs) were determined using the Bioconductor DESeq2 R-package with a false discovery rate of 0.05 and a minimum fold change value of 2 (Love, Huber and Anders, 2014). DEGs were determined as genes that displayed >1 or <-1 $\log_2$ fold change from pathogen treated samples compared to non-infected samples of the infection time point. The expression profile of each DEG replicate was then analysed by principle component analysis and hierarchical cluster analysis in MeV V 4.9 to show that the replicates for each time point and treatment were highly similar to one another and different from DEGs at other timepoints and treatments (Howe et al., 2011). Once the quality of the data had been determined for each replicate, the average normalised expression values for the three replicates of each DEGs at each time point was then transformed to Z-score values and a K-means cluster analysis was performed on the Z-score data of genes showing differential expression at least one time point. The K-means cluster analysis divided the DEGs into 10 clusters based on their expression profiles using the iDEP 9.0 K-means cluster function based on linkage averages and Pearson correlation coefficients between the different gene expression profiles (Ge, Son and Yao, 2018).

Transcription factor enrichment analysis was performed on the resulting clusters of DEGs in Pscan software of target gene promotor sequences 500 bp upstream of the start codon using transcription factor binding profiles from JASPAR 2018 framework (Zambelli, Pesole and Pavesi, 2009; Khan et al., 2018). P-values were generated for transcription factors whose motifs were enriched in the input clusters of potentially co-regulated genes identified using the K-means cluster analysis and transcription factor binding logos were obtained from JASPAR 2018 (Khan et al., 2018).

Phylogenetic analysis – sequence alignments and tree building

Phylogenetic analysis of the DA1 protein family was performed in plants that have full genome sequences available. BLASTP searches in NCBI (National Centre for Biotechnology Information) and BRAD (The Brassica Database) identified homologues of the *A. thaliana* DA1 and DAR proteins in the Brassicaceae family. Similar BLASTP searches in NCBI, Genome Database for Rosaceae (GDR), Legume Information service (LIS), Sol Genomics Network (SGN) and Phytozome using the full *A. thaliana* WRR5B and WRR7 amino acid sequences identified NLRs containing LIM or PEP domains from the annotated genomes of the plant kingdom.

Maximum likelihood (ML) trees of the Brassicaceae and *A. thaliana* sequences were generated using amino acid sequences starting from the LIM domain to the end of the protein sequence as predicted by HMMR or SMART (Schultz et al., 1998; Finn et al., 2015; Letunic, Ivica, Doerks and Bork, 2015). The ML tree of all the NLRs identified with either LIM or PEP domains was performed using full protein sequences. Sequence alignments were performed using MUSCLE software (Edgar, 2004), alignments were uploaded into SeaView or Jalview for sequence analysis (Gouy, Guindon and Gascuel, 2009; Waterhouse et al., 2009) and then used for maximum likelihood analysis using PhyML 3.0 algorithm or RAxML version 8.0 (Guindon et al., 2010; Stamatakis, 2014). All maximum likelihood analyses were performed using the LG model of evolution with 100 bootstrap replicates, all other parameters were left as the default. Resulting trees were annotated in iTOL v.5.4 (Letunic, I. and Bork, 2016). Phylograms showing the general evolutionary pattern of plant species were generated using PhyloT and based of NCBI genome data and visualised in iTOL v5.4 (Letunic, I. and Bork, 2016).

Synteny analyses

Synteny analysis was performed using CoGe SynMap2 software (Lyons and Freeling, 2008; Haug-Baltzell et al., 2017). Genomes of interest were aligned and a global synteny map was generated. Syntenic regions were then identified and target *A. thaliana* genes of interest were selected. Micro synteny output data was visualised in CoGe's GEvo tool where syntenic regions between the two genomes were joined by red markers (Lyons and Freeling, 2008).

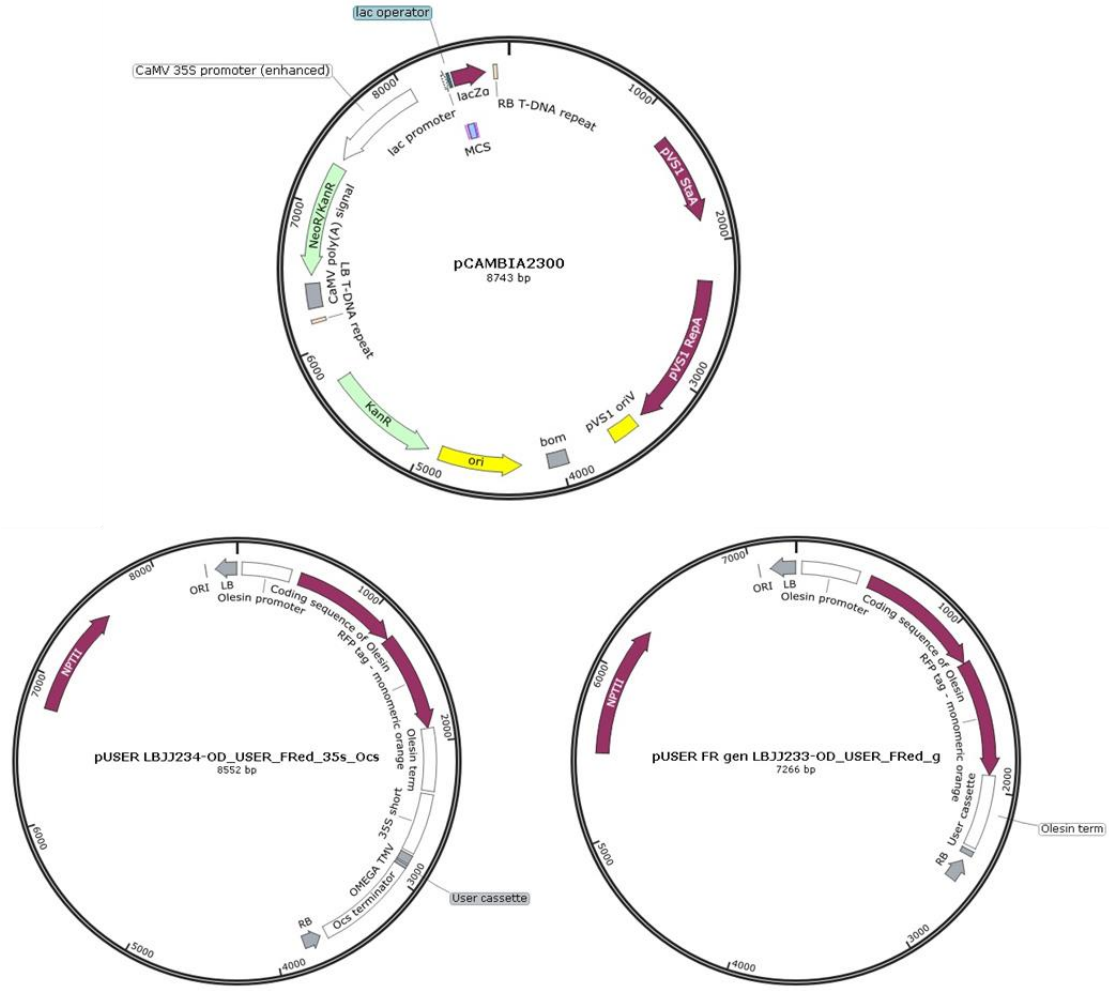


Figure 2.1: Vector maps
Vector maps of pCAMBIA2300 used for cloning of *WRR5A* and *WRR5B* as well as pUSER LBJJ233 and pUSER LBJJ234 that were used for USER cloning of genomic clones under the control of their native promoter and terminators or overexpression clones driven by the Cauliflower Mosaic Virus 35S promoter.

Table 2.1: List of Primers

Table of primers used for cloning (USER), genotyping T-DNA mutants and generating mutations (mutant). USER primers with overlapping sequences for specific vectors are shown in the 'comments' column (USER primers for LBJJ234 were for overexpression and therefore start at the start codon, USER primers for LBJJ233 were for genomic constructs and forward primers started in the gene of interests promotor region and reverse primers include the stop codon). Any reverse primer annotated 'for C-term tagging' has the stop codon removed.

Gene	Primer type	Direction	Sequence	Comments
WRR5A and B Overlap	USER	Forward	ACAGTTGCGUGGCT TCCGCTGGGCCATT TCAC	
WRR5A and B Overlap	USER	Reverse	ACGCAACTGUCATG ATACATCCCAACAA	
WRR5A	USER	Forward	GGCTTAAUATGACA AGCTCCTCCTCCTG GGT	For LBJJ234
WRR5A	USER	Reverse	GGTTTAAUTTAACA CAAAAGAGTGGAA CCAAAACCAG	For LBJJ234
WRR5A	USER	Reverse	AACCCGAUCCACA CAAAAGAGTGGAA CCAAAACCAG	For LBJJ234 and C-term tagging
WRR5B	USER	Forward	GGCTTAAUATGGA ACCACCAGCTGCTC GTG	For LBJJ234
WRR5B	USER	Reverse	GGTTTAAUTCATAA CTTTGAATATTGTG GAGTC	For LBJJ234
WRR5B	USER	Reverse	AACCCGAUCCTAAC TTTGAATATTGTGG AGTC	For LBJJ234 and C-term tagging
WRR5A-K239L	Mutant	Forward	GGATGCCCGGTATA GGTCTAACCACTCT CGCTACGATGC	P-loop mutant
WRR5A-K239L	Mutant	Reverse	GCATCGTAGCGAG AGTGGTTAGACCTA TACCGGGCATCC	P-loop mutant
WRR5B-K202L	Mutant	Forward	GTATGCCTGGCATA GGCCTGACGACGCT TGCTAAAGCAG	P-loop mutant
WRR5B-K202L	Mutant	Reverse	CTGCTTTAGCAAGC GTCGTCAGGCCTAT GCCAGGCATAC	P-loop mutant
RPS4	USER	Forward	GGCTTAAUATGGA GACATCATCTATTT C	For LBJJ234

RPS4	USER	Reverse	AACCCGAUCCGAA ATTCTTAACCGTGT	For LBJJ234 and C-term tagging
WRR7	Genotyping	Forward	GCATGGCAAGCGT GAGTACGAA	For RT-qPCR
WRR7	Genotyping	Reverse	TGCTCGAGTAACTT GTGTCCGATGA	For RT-qPCR
PP2AA3	Genotyping	Forward	GTTGTGGAGAACAT GATACGG	For RT-qPCR
PP2AA3	Genotyping	Reverse	GCTAGACATCATCA CATTGTC	For RT-qPCR
WRR7 promotor	Genotyping	Forward	AGGCAGTGGTACGT ACGTAC	
WRR7	USER	Forward	GGCTTAAUATGCCG ATCTCGGATGTCGC TTC	For LBJJ234
WRR7-C442Y	USER	Reverse	ACTGCAGCAUTTGG GAGTTCCATCAGC	For LBJJ234
WRR7-C442Y	USER	Forward	ATGCTGCAGUTACG AAAGGTTAGAGGT G	For LBJJ233
WRR7	USER	Reverse	GGTTTAAUTTACCT CCGGCGAAGAATCT CCTTG	For LBJJ233
WRR7	USER	Forward	AACCCGAUCCCCTC CGGCGAAGAATCTC CTTG	For LBJJ234 and C-term tagging
WRR7	USER	Forward	ATCGGGTUCGATGC CGATCTCGGATGTC GCTTC	
WRR7-S8F	USER	Forward	ATCTCGGATGTCGC TTTTTTGGTTGGAG GTGC	S8F mutation
WRR7	USER	Reverse	GGTTTAAUTTACCT CCGGCGAAGAATCT CCTTG	
DA1	USER	Forward	GGCTTAAUATGGGT TGGTTTAACAAGAT CTTTAAAGGC	For LBJJ234
DA1	USER	Reverse	AACCCGAUCCAAC CGGGAATCTACCGG TCATCTG	For LBJJ234 and C-term tagging
DAR1	USER	Forward	GGCTTAAUATGGG GTGGCTAACTAAAA TCCTTAAAGGTTC	For LBJJ235
DAR1	USER	Reverse	AACCCGAUCCAGG AAATGTACCGGTCA AGCGAATATGG	For LBJJ234 and C-term tagging

DAR2	USER	Forward	GGCTTAAUATGGATTCTTCTTCCTCTTCCTCTTCTTCTTC	For LBJJ236
DAR2	USER	Reverse	AACCCGAUCCCAAAGGAAAAGTTCCAGTTAAGCGG	For LBJJ234 and C-term tagging
DAR3	USER	Forward	GGCTTAAUATGTCGTGTTGCTTCTCCTGCTTCAA	For LBJJ237
DAR3	USER	Reverse	AACCCGAUCCAACCGTTGAATCTGGTGTAGCGTTG	For LBJJ234 and C-term tagging
DAR6	USER	Forward	GGCTTAAUATGGCCTCTGATTACTATTCATCTGAC	For LBJJ238
DAR6	USER	Reverse	AACCCGAUCCCCTCTGGCGAAGAATCTCTTTCAGG	For LBJJ234 and C-term tagging
DAR7	USER	Forward	GGCTTAAUATGTGGTGTTTGTCCTGCTTTAAACCT	For LBJJ239
DAR7	USER	Reverse	AACCCGAUCCAAGATTTGAATCTGGTTTCGTCCAG	For LBJJ234 and C-term tagging
EOD1	USER	Forward	GGCTTAAUATGAATGGAGATAATAGACCAGTGG	For LBJJ240
EOD1	USER	Reverse	AACCCGAUCCATGAATGCTGGGCTCCCCAAAGAC	For LBJJ234 and C-term tagging
DA1 LIM-Peptidase	USER	Forward	ATGTGTGCUGGCTGTAATATGGAG	For WRR7 fusion
DA1 LIM-Peptidase	USER	Reverse	AGCACACAUGCTACAAGTCAAAAGATACT	For WRR7 fusion
DAR1 LIM-Peptidase	USER	Forward	AGCATGTGUGTCGGTTGCCAAGC	For WRR7 fusion
DAR1, DAR6 and DAR7 LIM-Peptidase	USER	Reverse	ACACATGCUACAAGTCAAAAGATACT	For WRR7 fusion
DAR3 LIM-Peptidase	USER	Forward	AGCATGGAUGGCAAATCTGAGATTGG	For WRR7 fusion
DAR3 LIM-Peptidase	USER	Reverse	ATCCATGCUACAAGTCAAAAGATACT	For WRR7 fusion
WRR5B LIM-Peptidase	USER	Forward	AGCATGUGCAAGGATTGCAAATCTGCA	For WRR7 fusion

WRR5B LIM-Peptidase	USER	Reverse	ACATGCUACAAGTCAAAAGATACT	For WRR7 fusion
DAR6 LIM-Peptidase	USER	Forward	AGCATGTGUGGTGGTTGCAACTTTGC	For WRR7 fusion
DAR7 LIM-Peptidase	USER	Forward	AGCATGTGUGATGGTTGCAAATCTGC	For WRR7 fusion
DA1STP::WRR7Term	USER	Forward	ATCGTCTUTAAACCGGGAATCTACCGGTC	For WRR7 fusion
DAR1STP::WRR7Term	USER	Forward	ATCGTCTUTAAGGAAATGTACCGGTCAAGC	For WRR7 fusion
DAR3STP::WRR7Term	USER	Forward	ATCGTCTUCAAACCGTTGAATCTGGTGTAGC	For WRR7 fusion
WRR5BSTP::WRR7Term	USER	Forward	ATCGTCTUCATAACTTTGAATATTGTGGAGTC	For WRR7 fusion
DAR6STP::WRR7Term	USER	Forward	ATCGTCTUTACCTCTGGCGAAGAATCTCTTTCAG	For WRR7 fusion
DAR7STP::WRR7Term	USER	Forward	ATCGTCTUCAAAGATTTGAATCTGGTTTCGTC	For WRR7 fusion
WRR5B LIM-Peptidase	USER	Forward	ATAGGATUTGCAAGGATTGCAAATCTGCA	For DA1 fusion
WRR7 LIM-Peptidase	USER	Forward	ATAGGATUTGTGGTGGTTGCAACTC	For DA1 fusion
DAR1 LIM-Peptidase	USER	Forward	ATAGGATUTGTGTCGGTTGCCAAGCTG	For DA1 fusion
WRR5B, WRR7 and DAR1 LIM-Peptidase	USER	Reverse	AATCCTAUGCAGACAAGTTGTTGACAAG	For DA1 fusion
CAMTA2	USER	Forward	GGCTTAAUGAGTTGACTCATCTTACGACAC	For LBJJ233
CAMTA2	USER	Reverse	GGTTTAAUCTACAGCAGCTGATGACAGCTTC	For LBJJ233
CAMTA2	USER	Forward	GGCTTAAUATGGCGGATCGCGGATCTTTCGGA	For LBJJ234
CAMTA2	USER	Reverse	AACCCGAUCCTTCAAATGCAAGAGACATGAAAGTGT	For LBJJ234 and C-term tagging

CAMTA3	USER	Forward	GGCTTAAUATGGCG GAAGCAAGACGAT TCAG	For LBJJ234
CAMTA3	USER	Reverse	AACCCGAUCCACTG GTCCACAAAGATG AGGACATAG	For LBJJ234 and C-term tagging
CAMTA1	USER	Forward	GGCTTAAUATGGTG GATCGCAGATCTTT TGGC	For LBJJ234
CAMTA1	USER	Reverse	AACCCGAUCCAGG AGAAATAGACATC ATCAATGTGTCATC	For LBJJ234 and C-term tagging
CAMTA2	USER	Forward	ATCGGGTUCGATGG CGGATCGCGGATCT TTCGGA	For N term Tagging-F
CAMTA2	USER	Reverse	GGTTTAAUTCATTC AAATGCAAGAGAC ATGAAAGTG	With Stop codon
CAMTA2	Genotyping	Forward	GGAACCTCCACTTC TCCAAAC	SALK_007027
CAMTA2	Genotyping	Reverse	CCCTGTTAACGTCA GAGCATC	SALK_007027
CAMTA1	Genotyping	Forward	CAGGTTCCATGATT GGAAAAC	SALK_008187
CAMTA1	Genotyping	Reverse	ACTCAGATCGGTTA GGGTTCG	SALK_008187
CAMTA3	Genotyping	Forward	TGAAAACCTGATGA ATCCGAG	SALK_001152
CAMTA3	Genotyping	Reverse	GGTTGTGAAGTGGT GGTAAGC	SALK_001152
Salk T-DNA insert	Genotyping	Reverse	ATTTTGCCGATTTC GGAAC	SALK genotyping
BAK1	USER	Forward	GGCTTAAUATGGA ACGAAGATTAATG ATCCCTTG	For LBJJ234
BAK1	USER	Reverse	AACCCGAUCCTCTT GGACCCGAGGGGT ATTCG	For LBJJ234 and C-term tagging
MAP3Kδ4	USER	Forward	GGCTTAAUTCGCCA TTAGAAAAGAGAA AGCT	For LBJJ233
MAP3Kδ4	USER	Reverse	GGTTTAAUGATCCA CCAACACAAGCGA T	For LBJJ233
proCHR4	USER	Forward	GGCTTAAUCCAGTG CAGAACACGACGT GGAGTG	For LBJJ233 Fragment 1
CHR4	USER	Reverse	AGAATGTCUACAA AATCTATCATGAG	For LBJJ233 Fragment 1

CHR4	USER	Forward	AGACATTCUACTATGCTGATACTAG	For LBJJ233 Fragment 2
CHR4	USER	Reverse	AGGCTGCUGAAAAGGTTCTTCCTT	For LBJJ233 Fragment 2
CHR4	USER	Forward	AGCAGCCUCTTAGTAATATGGATGG	For LBJJ233 Fragment 3
CHR4	USER	Reverse	GGTTTAAUGATGTAGCCTAGTCTGATCCCGAG	For LBJJ233 Fragment 3
MAC7	USER	Reverse	GGTTTAAUCGTTCTTTCGTTTCTTTGTCAGT	For LBJJ233
MAC7	USER	Forward	GGCTTAAUTGGACTCGTGGATGCAACATCAC	For LBJJ233
CaM2	USER	Forward	GGCTTAAUATGGCGGATCAGCTCACAGACGATC	For LBJJ234
CaM2	USER	Reverse	AACCCGAUCCCTTAGCCATCATAACCTTCACAAACTC	For LBJJ234 and C-term tagging
CaM3	USER	Forward	GGCTTAAUATGGCGGATCAGCTCACCGAC	For LBJJ234
CaM3	USER	Reverse	AACCCGAUCCCTTAGCCATCATGACCTTAAC	For LBJJ234 and C-term tagging
YFP C termTag	USER	Forward	ATCGGGTUCGATGGTGAGCAAGGGCGAGGAG	
YFP C termTag	USER	Reverse	GGTTTAAUAAGCTCACTTGTACAGCTCGT	For LBJJ234
YFP N term Tag	USER	Forward	GGCTTAAUATGGTGAGCAAGGGCGAGGAG	For LBJJ234
YFP N term Tag	USER	Reverse	AACCCGAUCCCTTGTACAGCTCGTCCAT	
V5 N term Tag	USER	Forward	GGCTTAAUATGCATTCGGGTAAGCCAATC	For LBJJ234
V5 N term Tag	USER	Reverse	AACCCGAUCCGGTTGAGTCGAGTCCGAGC	
HF C term Tag	USER	Forward	ATCGGGTUCGCATTCGGGTTCCGGAAGAGGA	

HF C term Tag	USER	Reverse	GGTTTAAUAAGCTC ACTTGTCATCGTCA	For LBJJ234
V5 C term Tag	USER	Forward	ATCGGGTUCGCATT CGGGTAAGCCAATC CC	
V5 C term Tag	USER	Reverse	GGTTTAAUAAGCTT AGGTTGAGTCGAGT CCGAG	For LBJJ234
GFP C term Tag	USER	Forward	ATCGGGTUCGAAA GAGTTCATGCGCTT C	
GFP C term Tag	USER	Reverse	GGTTTAAUTCAGCG GCCCTCGGAGCGC	For LBJJ234

Chapter 3

Identification of multiple resistance genes conferring immunity to *Albugo candida* isolate AcEM2 in the *Arabidopsis thaliana* accession Col-0

Introduction

Pathogens and hosts are continually embroiled in an evolutionary arms race of invasion, detection and evasion. In plants, this has led to a multi-layered phytopathogen detection system comprised of cell surface and intracellular receptors that induce defence responses (Jones, J., D. G. and Dangl, 2006; Dodds and Rathjen, 2010). Pattern recognition receptors (PRRs) are located at the cell surface and function by detecting conserved microbe associated molecular patterns (MAMPs) such as flagellin or chitin leading to MAMP triggered immunity (MTI). NB-ARC Leucine Rich Repeat receptors (NLRs) are intracellular immune receptors that detect the intracellular presence of pathogen-derived proteins (effectors) that impose susceptibility, often by supressing MTI. Once NLRs detect a pathogens presence they elicit effector triggered immunity (ETI), resulting in the salicylic acid (SA) induced hypersensitive response (HR) killing the infected cell (Dodds and Rathjen, 2010).

Over recent years, we have gained an increasing understanding of the mechanistic function of NLRs and the diversity of ways that plants utilise NLRs to detect pathogen effectors. These include direct interaction of NLR and effector or sensing the effectors presence via its effect on an intermediary immune associated protein known as a guardee (if it retains its normal cellular function) or a mimic of the effectors target known as a decoy that has lost its original host function (Jones, J.D.G., Vance and Dangl, 2016; Cesari, 2017). NLRs fall into two discrete sub-classes (TNLs or non-TNLs) based on the presence or absence of a Toll-interleukin receptor (TIR) at their N-terminus. Non-TNLs are often referred to as CNLs because a large number of them contain a N-terminal coiled coil motif and more recently a sister group to CNLs has been identified containing a Resistance to Powdery Mildew 8 (RPW8) domain at their N-terminus and are referred to as CC_Rs (Shao et al., 2016;

Zhang, Y.-M. et al., 2016; Nepal et al., 2017). The distinction of NLR sub-classes is important because TNL, CNLs and CC_Rs induce the production of SA via different pathways (Bonardi et al., 2011). TNLs require a lipase like protein Enhanced Disease Susceptibility 1 (EDS1) which heterodimerises with either Phytoalexin deficient 4 (PAD4) or Senescence-associated carboxylesterase 101 (SAG101) to induce defence, whereas CNLs and CC_Rs act independently of EDS1 (Wiermer, Feys and Parker, 2005; Cui et al., 2017; Lapin et al., 2019). Our understanding of CNL and CC_R downstream signalling is poor, although the ETI signalling of some CNLs has been linked with plasma membrane bound Non Race specific disease resistance 1 (NDR1) (Day, Dahlbeck and Staskawicz, 2006b; McNeece et al., 2017).

NLRs of all classes contain a NB-ARC domain, that is part of the signal transduction ATPases with numerous domains (STAND) that contain a P-loop motif that binds ATP and ADP (Leipe, Koonin and Aravind, 2004). The presence of ATP or ADP at this site determines whether the NLR is active, recent analysis of CNL HopZ-activated disease resistance 1 (ZAR1) has shown that it is inactive when bound to ADP and that ADP is released following the recognition of the AvrAC effector from bacterial phytopathogen *Xanthomonas campestris* pv. *campestris* (Wang, J., Wang, et al., 2019). The P-loop is made up of the Walker A and Walker B motifs that form the phosphate binding site that is capable of binding to ATP an ADP, the Walker A motif is characterised by a GxxxxGK sequence (Leipe, Koonin and Aravind, 2004). The binding of ADP and ATP to the P-loop can be disrupted by mutating the conserved positively charged lysine residue (Tameling et al., 2006; Slootweg et al., 2010; Sohn et al., 2014). Therefore, P-loop mutants can be used to test whether particular NLRs are required for defence activation.

Originally, the gene-for-gene hypothesis proposed that one plant resistance gene would recognise one avirulence gene (Flor, 1971). However, work over the last few decades has determined NLRs and NLR complexes that recognise effectors from multiple pathogens (Narusaka et al., 2009; Bonardi et al., 2011; Wu, C.-H. et al., 2017). In addition, several NLRs have been found to act as obligate pairs, including RPS4/RRS1, RGA4/RGA5 and Pikp-1/Pikp-2 (Narusaka et al., 2009; Eitas and Dangl, 2010; Césari et al., 2014; Le Roux et al., 2015; Maqbool et al., 2015; Sarris, Panagiotis F. et al., 2015; Huh et al., 2017). Each pair contains one NLR that's role is to recognise the presence of the pathogen and is known as the 'sensor' NLR and

the other executes immune signalling and is known as the 'helper' or 'executioner' NLR. So far, in dual systems only the executioner NLR requires motifs that are essential for NLR activation such as the P-loop and the helper NLR often loses some of these traditionally conserved motifs (Sohn et al., 2014). Many dual NLR detection systems have been shown to act through the direct binding of pathogen effectors to additional non-canonical domains in the sensor NLR e.g. WRKY in RRS1 or the Heavy metal associated (HMA) domain of RGA5 and Pikp-1, followed by the execution of defence signalling by its partner NLR (Cesari et al., 2013; Le Roux et al., 2015; Maqbool et al., 2015). Although, more recent studies have found that non-canonical, integrated domain containing NLRs can detect effectors through indirect physical association with the effectors target protein e.g. the NOI integrated domain of OsPii-2 binds to Os-Exo70-F3 in a manner akin to the guard model (Jones, J.D.G., Vance and Dangl, 2016; Fujisaki et al., 2017). Our understanding of the role integrated domains play in NLR mechanics is in its infancy. Currently, over 265 NLR fusions have been determined from 40 genomes, including bryophytes, monocots and dicots (Kroj et al., 2016; Sarris, P. F. et al., 2016) and in the *Arabidopsis thaliana* pan-NLRome alone there are 36 known integrated domains (Van de Weyer et al., 2019). However, only a few of these have been studied in any depth. Studying more of these systems will increase the arsenal of NLRs and NLR derivatives that can be deployed to combat crop diseases.

Albugo candida is an obligate biotrophic Oomycete pathogen that causes White blister rust disease on over 200 Brassicaceae species, including important vegetable and oilseed crops as well as the model plant *A. thaliana*. *A. candida* is a strong immunosuppressor and therefore not only causes primary infection yield losses but also exposes crops to secondary infection by non-host pathogens such as *Phytophthora infestans* and *Hyaloperonospora arabidopsis* (Cooper et al., 2008; Prince et al., 2017). Although, *A. candida* can infect over 200 Brassicaceae species, it has evolved distinct physiological races that have developed specialisms to individual species (Borhan, M. Hossein et al., 2008; McMullan et al., 2015). Usefully, *A. thaliana* ecotypes display differing levels of susceptibility to *A. candida* races e.g. Col-0 is completely resistant to race 4 isolate AcEM2 whereas Ws-2 is fully susceptible (Borhan, M. Hossein et al., 2008). This diversity of defence response is due to underlying genetic factors that are present or absent from different

A. thaliana ecotypes. Therefore, we can use the genetic diversity of *A. thaliana* ecotypes to identify novel resistance genes that are active against *A. candida* by generating recombinant inbred lines derived from crosses of resistant and susceptible ecotypes followed by phenotype-based mapping.

Here we identify multiple resistance genes active against *A. candida* isolate AcEM2 in the Col-0 background. These include the previously characterised *A. candida* resistance gene *WRR4A* (Borhan, M. Hossein et al., 2008), novel resistance gene pair *WRR5A* and *WRR5B* as well as an atypical resistance protein encoding gene *WRR7*. *WRR5A* (*CSA1*) and *WRR5B* (*CHS3*) were previously shown to cause autoimmune responses but up to this point had no known associated resistance function against a plant pathogen (Xu et al., 2015). *WRR7* encodes an NLR-like protein that contains an N-terminal RPW8 domain a partial NB-ARC domain and a C-terminal integrated LIM-peptidase domain but lacks a Leucine rich repeat (LRR).

Two of these proteins, WRR5B and WRR7, contain the same additional non-canonical integrated domains: A LIN_{11}, ISL_1 and MEC_3 (LIM) domain and a zinc metallopeptidase (Peptidase) domain (Dong, H. et al., 2017). The tandem arrangement of the LIM and peptidase domains are found in proteins exclusive to the plant kingdom and the presence of these two domains demark them as DA1 protein family members. In *A. thaliana,* this protein family is comprised of eight proteins, DA1 and DA1-related (DAR) 1-7 (Srivastava and Verma, 2017). WRR5B (DAR4) and WRR7 (DAR5) are the only two resistance genes that are part of this protein family in *A. thaliana*, the rest either function in the regulation of cell size or have no known function (Peng, Yuancheng et al., 2015). This hints at a common mechanism of recognition that could be shared between these resistance pathways.

Results

Genetic analysis reveals multiple independent *White Rust Resistance* genes in the *Arabidopsis thaliana* Columbia accession

Arabidopsis thaliana ecotypes Col-0 and Ws-2 respond differently to infection by *A. candida* isolate AcEM2. The Columbia ecotype is fully resistant, whereas the Ws-2 ecotype is susceptible. One resistance gene, *White Rust Resistant 4A* (*WRR4A*) is already known to be active against *A. candida* isolate AcEM2 in the Col-0 background (Borhan, M. Hossein et al., 2008). However, analysis of recombinant inbred lines (RILs) that were derived from an original Col-5 x Ws-2 cross and subsequently selfed for eight generations (Eric Holub, University of Warwick) revealed the presence of multiple *WRR* genes in the Columbia genome conferring distinct resistance responses against *A. candida* isolate AcEM2 (Eric Holub, personal communication) (Fig 3.1). To determine whether other genetic elements were responsible for the different resistance responses observed in the Col-5 x Ws-2 RILs (termed CW RILs), seedlings were scored from 0 (fully resistant with no visible necrotic lesions) to 5 (fully susceptible with no cell death response) (Eric Holub, personal communication). A Quantitative trait loci (QTL) analysis was then performed on these populations to identify additional resistance gene loci (Eric Holub and Volkan Cevik, personal communication).

The QTL analysis revealed three loci that were associated with the resistance phenotypes observed (Fig 3.2). The *WRR4A* locus was responsible for the strongest resistance phenotype, where no visible cell death was observed but the weaker resistance phenotypes were associated with two regions on chromosome 5. These two loci were located towards the telomeric ends of chromosome 5, on both chromosome arms, showing that two distinct regions were involved in the cell death responses observed in the RILs.

Fine mapping of the *WRR5* locus was then carried out, revealing that two resistance genes were present in this locus. These were *CSA1* (named as *WRR5A*) encoding a canonical TNL protein and *CHS3 (At5g17890)* (named as *WRR5B*) encoding TNL with an integrated LIM and zinc metallopeptidase (Peptidase) domain (Dong, H. et al., 2017). In the *WRR7* QTL region, we identified three resistance gene candidates. These included the canonical CC_R type of resistance protein encoding genes *NRG1.1* (*At5g66900*) and *NRG1.2* (*At5g66910*) as well as a gene encoding an atypical

resistance protein, *DAR5* (*At5g66630*) which encodes a CC_R resistance protein containing an integrated LIM and Peptidase integrated domain.

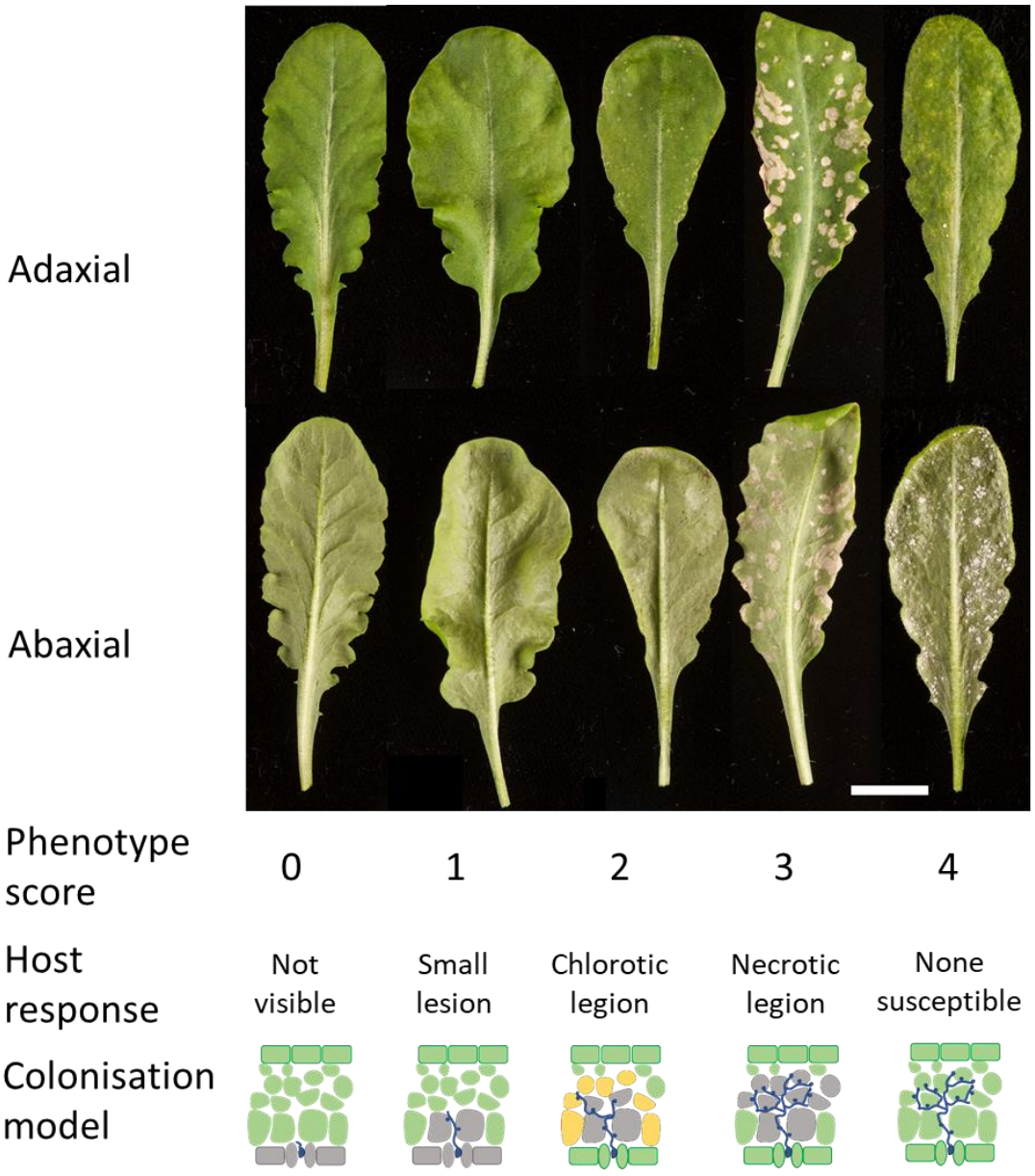
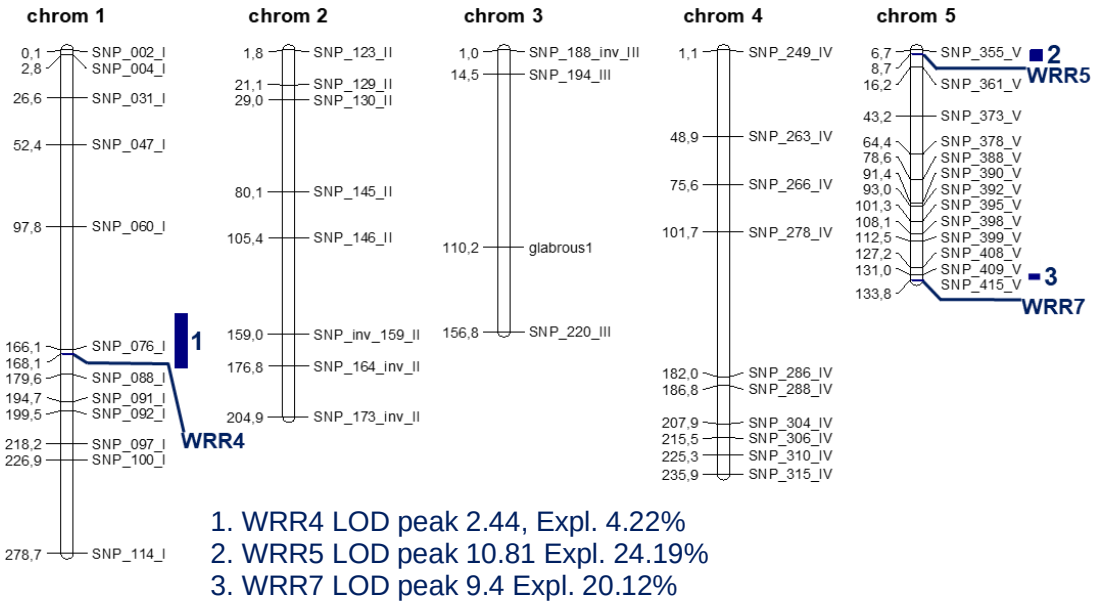


Figure 3.1: Post-infection phenotypes of Col-5 x Ws-2 recombinant inbred lines inoculated with *Albugo candida* isolate AcEM2
Representative adult leaf phenotypes of Col-5 x Ws-2 Recombinant inbred lines (RIL) 7 days post infection with *Albugo candida* isolate AcEM2 and a diagrammatic depiction of the level of AcEM2 colonisation, grey and yellow cells represent cells undergoing a resistant cell death response, blue represents *A. candida* hyphae. Scale bar represents 1cm. RIL lines were generated and scored by Eric Holub and Volkan Cevik.



1. WRR4 LOD peak 2.44, Expl. 4.22%
2. WRR5 LOD peak 10.81 Expl. 24.19%
3. WRR7 LOD peak 9.4 Expl. 20.12%

Figure 3.2: QTL map of *WRR* loci associated with Col-5 x Ws-2 recombinant inbred line resistance phenotypes

Arabidopsis thaliana chromosome map from Cevik and Holub (2016) showing QTL positions and SNP markers based on 83 recombinant inbred lines from a Col-5 x Ws-2 cross. Distances (cM) were estimated using Haldane's mapping function, blue bars represent identified QTL regions at a 5% significance level that are associated with resistance traits to *Albugo candida* isolate AcEM2. Logarithm of odds (LOD) scores and the percentage of variance explained by the QTL (Expl) for each QTL region are shown in at the base of the map. QTL 1 is large due to a second peak identified on chromosome 1 with a LOD score below the 10% threshold and the region is extended to include this peak.

Characterization of distinct resistance phenotypes mediated by different *White Rust Resistance* genes in *Arabidopsis thaliana* accession Columbia

To determine the distinct resistance phenotypes mediated by different resistance loci that were identified in the QTL analysis, we identified RILs containing just one of the three identified *WRR* loci derived from the resistant Col-5 progenitor. The RILs identified were CW20 ($WRR4A^{Col\text{-}0}/wrr5^{Ws\text{-}2}/wrr7^{Ws\text{-}2}$), CW5 ($wrr4a^{Ws\text{-}2}/WRR5^{Col\text{-}0}/wrr7^{Ws\text{-}2}$), CW14 ($WRR4^{Col\text{-}0}/wrr5^{Ws\text{-}2}/WRR7^{Col\text{-}0}$) and fully susceptible line CW234 ($wrr4a^{Ws\text{-}2}/wrr5^{Ws\text{-}2}/wrr7^{Ws\text{-}2}$). The RILs displayed four distinct phenotypes following pathogen infection. Both parental phenotypes were observed in the RILs: a green resistance phenotype was observed with the RIL CW20 that contained functional $WRR4A^{Col\text{-}0}$ only, similar to the phenotype as observed in the resistant Col-5 progenitor ecotype. Two additional resistant phenotypes observed were the chlorotic resistance response with the RIL CW5 containing $WRR5^{Col\text{-}0}$ and the necrotic resistance response observed in the RIL CW14 with $WRR7^{Col\text{-}0}$ (Fig 3.3). This analysis revealed that these resistance gene loci confer distinct resistance phenotypes independently and there does not seem to be additivity.

Trypan blue staining of the RIL lines revealed that cell death occurred in all three resistant lines including the green resistant line (CW20) and Col-5 parent. However, the cell type undergoing cell death differed between the green resistant and chlorotic/necrotic resistant lines. Cell death observed in the Col-5 parent and green resistant line was predominantly confined to the epidermal cell layer whereas cell death in the chlorotic and necrotic resistant lines was much more extensive than the cell death observed in the green resistant and Col-5 ecotype. Intriguingly, we observed that the cell death of a single epidermal cell directly adjacent to the stomatal guard cell, where the initial point of pathogen entry was located, was enough to block pathogen growth in CW20 plants. This was in contrast to the CW5 and CW14 lines, where trypan blue staining revealed that cell death was extensive and spread to the mesophyll cells (Fig 3.3). Therefore, $WRR4A^{Col\text{-}0}$ operates in a different spatial location within the leaf to the resistance genes underpinning the resistance phenotypes observed in CW5 and CW14 lines which confer a weaker resistance response as observed with $WRR4A^{Col\text{-}0}$. The tissue specific operation of these resistance mechanisms leads to $WRR4A^{Col\text{-}0}$ being spatially epistatic over the genes underpinning chlorotic and necrotic resistance.

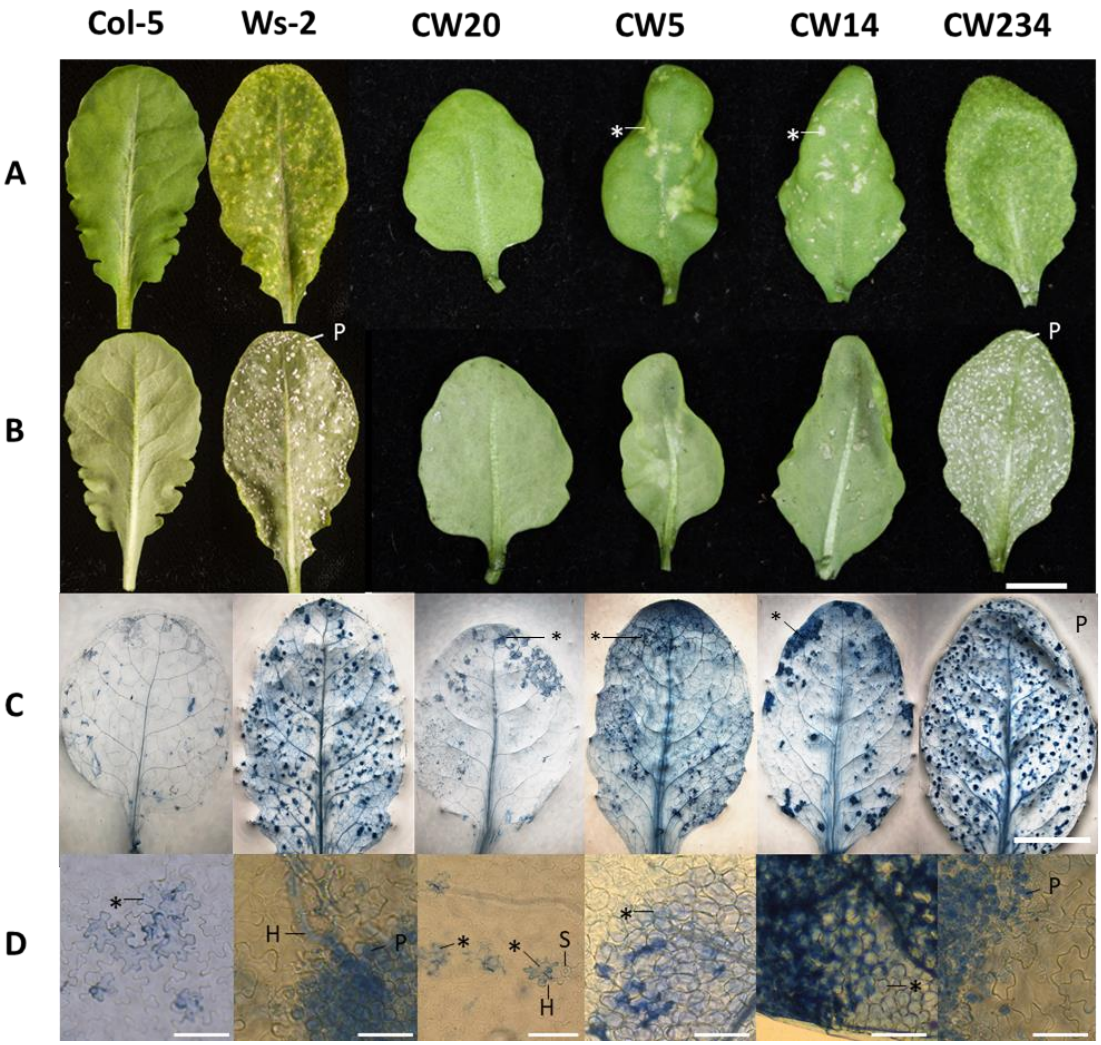


Figure 3.3: Phenotypes of single *WRR* loci containing recombinant inbred lines post infection by *Albugo candida* isolate AcEM2
Col-5 x Ws-2 recombinant inbred line (CW) phenotypes 7 days post infection with *Albugo candida* isolate AcEM2, tiles in row A show the adaxial leaf surface and tiles in row B show the abaxial surface. Images in rows C and D show representative leaves stained with Trypan Blue dye, staining patterns show pathogen structures within the infected leaf as well as dead host cells. Areas of cell death are indicated with an *, pustule formation is marked with P, hyphal growth and stomata are indicated with H and S respectively. Scale bars for rows A-C represent 1cm, scale bars in row D show 100µm.

Cloning of resistance genes in the *WRR5* locus

Fine mapping carried out by Eric Holub and Volkan Cevik identified two resistance gene candidates in the *WRR5* locus, *CSA1* (named as *WRR5A*) and *CHS3* (named as *WRR5B*). These two resistance genes are located adjacent to each other in the Col-0 genome, in a tandem orientation and sharing a promoter region. These two genes have previously been shown to be required for autoimmunity induced by the *chs3*-2D mutant (Xu et al., 2015). Therefore, we hypothesized that both *WRR5A* and *WRR5B* would be required for the resistance response to AcEM2. To test this hypothesis, genomic clones of both *WRR5A* and *WRR5B* were constructed individually and together. The single clone containing both genes and the promotor element spanned 20,272 bp, starting 2704 bp upstream of the *WRR5A* stop codon to 1424 bp downstream of the *WRR5B* start codon. Individual clones were also constructed with the promotor sequence of WRR5A 2704 bp upstream of the start codon and the terminator region extending 1031 bp past the stop codon, the WRR5B promotor region spanning 1213 bp upstream of the start codon and the terminator region encompassing 1424 bp downstream of the stop codon. These constructs were subsequently cloned into *pCAMBIA2300* vector and transformed into AcEM2 susceptible *A. thaliana* ecotype Ws-2. The resulting homozygous T_3 transgenic lines derived from independent transgenic events were then challenged with AcEM2 to identify whether the genomic clones were sufficient to elicit an immune response.

We found that Ws-2 transgenic lines containing individual *WRR5A* or *WRR5B* constructs were unable to provide resistance to *A. candida* isolate AcEM2. However, Ws-2 lines containing the genomic clone of both *WRR5A* and *WRR5B* were able to provide a chlorotic resistance response to AcEM2, similar to the phenotype observed in the RIL CW5 (Fig 3.5). Therefore, both *WRR5A* and *WRR5B* are required to confer the chlorotic resistance response in *A. thaliana* against *A. candida* isolate AcEM2.

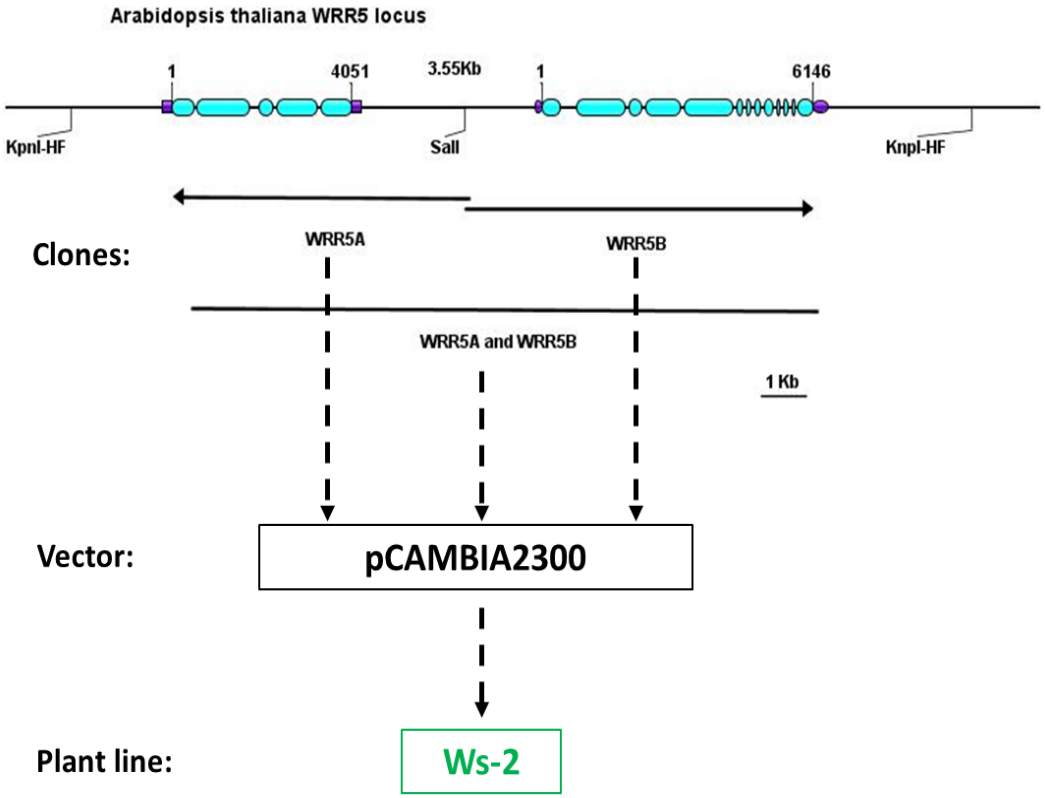


Figure 3.4: *WRR5A* and *WRR5B* genomic clones
Schematic diagram depicting the *WRR5* loci and the genomic clones made for complementation assay in *Arabidopsis thaliana* ecotype Ws-2. Genome data was visualised using the TAIR genome browser of the *A. thaliana* Col-0 genome (Lamesch et al., 2011). Blue rectangles show the exon positions, purple boxes show untranslated regions, the restriction enzyme cut sites are marked below the nucleotide diagram and upper numbers show the start and stop codon positions. Scale bar represents 1Kb.

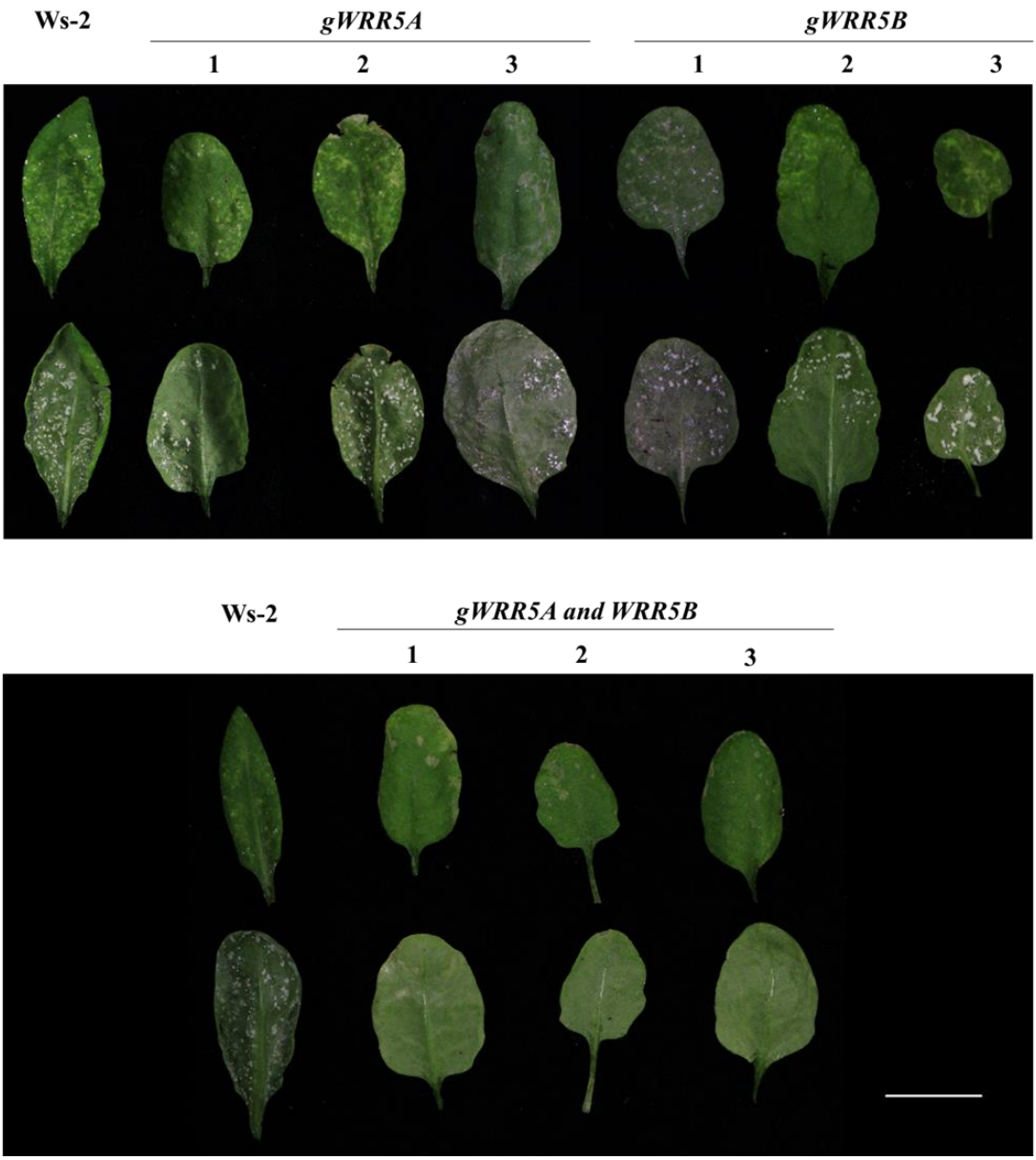


Figure 3.5: *WRR5A* and *WRR5B* compliment the AcEM2 susceptible phenotype of Ws-2 together but not individually
Post infection phenotypes of three independent homozygous T_3 resistant lines of *Albugo candida* susceptible *Arabidopsis thaliana* ecotype Ws-2 transformed with either *gWRR5A*, *gWRR5B* or *gWRR5A* and *gWRR5B*, 7 days post infection with *A. candida* isolate AcEM2. Scale bar represents 1cm.

An intact P-loop in both WRR5A and WRR5B is required to cause an immune response

Our transformation experiments revealed that both *WRR5A* and *WRR5B* are required for immune response against *A. candida* isolate AcEM2 (Fig 3.5). These two resistance genes are located in a tandem head to head orientation in the genome. *WRR5B* is the previously reported *CHS3* (*Chilling sensitive 3*) that contains a non-canonical integrated LIM-Peptidase domain (Yang et al., 2010). A mutant of *WRR5B* (*chs3*-2D) causes autoimmunity which is dependent on its tandem NLR partner *CSA1* (*WRR5A*) (Xu et al., 2015). Our finding that WRR5A and WRR5B are both required to stimulate an immune response in *A. thaliana* against *A. candida* isolate AcEM2, in combination with studies showing the requirement of WRR5A for the autoimmune phenotype of the *WRR5B chs3-2D* autoimmune allele, suggests that WRR5A and WRR5B act by the sensor-helper model of NLR activation to stimulate immunity. Previous studies of sensor-helper NLR pairs have shown that the NLR containing an integrated domain (in our case WRR5B) senses the presence of an effector and the canonical NLR (WRR5A) activates defence. Defence activation of NLRs is known to require the P-loop motif from the NB-ARC domain which binds ATP and ADP (Leipe, Koonin and Aravind, 2004; Wang, J., Wang, et al., 2019). In sensor-helper NLR activation such as in the RRS1-RPS4 and RGA4 and RGA5 systems, it has been shown that only an intact P-loop of the executioner NLR (RPS4/RGA4) is required for immune activation (Césari et al., 2014; Sohn et al., 2014). The P-loop is characterised by several conserved motifs, one of which is the GxxxxGK Walker A motif which is involved in the binding of ATP to the NB-ARC domains (Tameling et al., 2006). Previous studies have shown that mutation of the conserved lysine residue in the Walker A motif results in the loss of function of the P-loop (Tameling et al., 2006; Slootweg et al., 2010; Sohn et al., 2014). Therefore, to test whether WRR5A and WRR5B operate via the sensor-helper model akin to the system utilised by RRS1 and RPS4 we generated P-loop mutants of the conserved lysine residue in the Walker A motif of WRR5A (WRR5A-K239L) and WRR5B (WRR5B-K202L) as well as the autoimmune mutant of WRR5B (WRR5B-C1340Y) (Fig 3.6) and co-expressed these proteins in *N. tabacum* to determine whether combinations of these mutants could cause autoimmune responses.

We found that autoimmunity conferred by the WRR5B- C1340Y mutation could be induced following the co-expression with *WRR5A* construct in transient cell death experiments in *N. tabacum*. We also found that the autoimmune response induced by the WRR5B- C1340Y mutant could not be induced when co-expressed with WRR5A-K239L or when WRR5B- C1340Y itself had a compromised P-loop. This reveals that intact P-loops of both WRR5A and WRR5B are required for immune activation (Fig 3.7), in contrast to the mechanisms observed with RRS1 and RGA5 (Césari et al., 2014; Sohn et al., 2014).

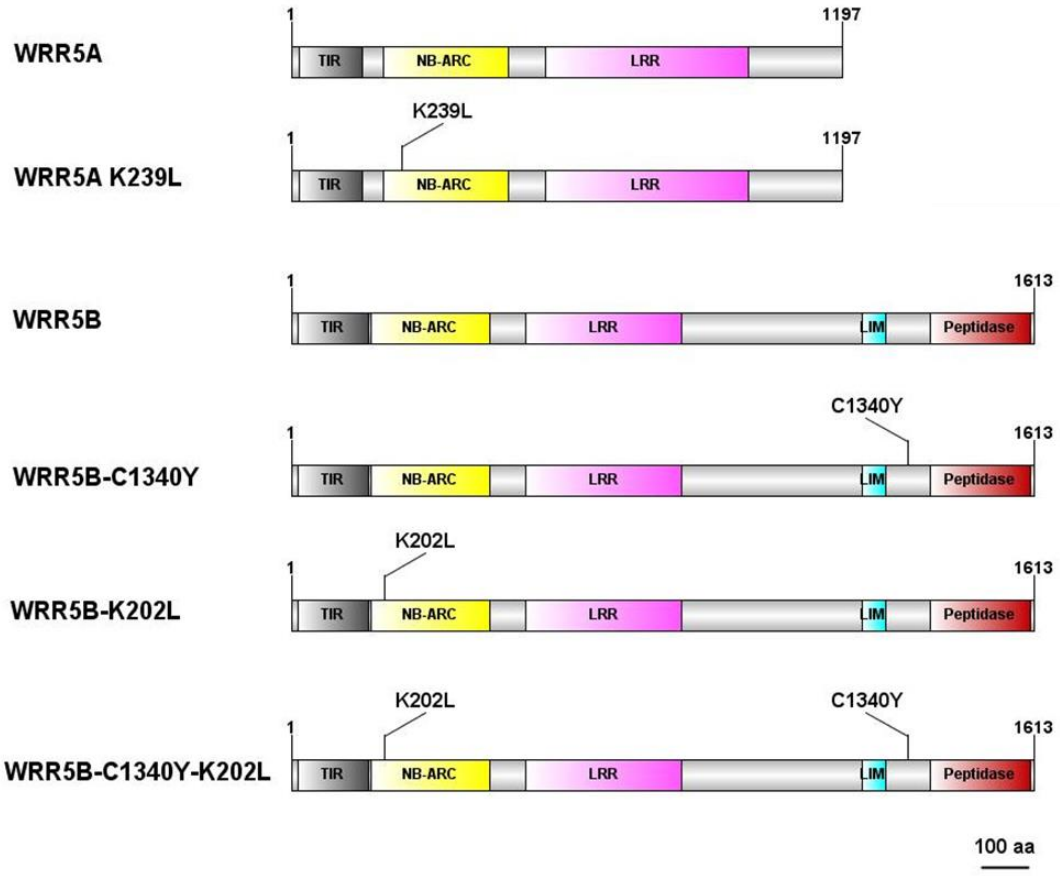


Figure 3.6: Cloning of WRR5A and WRR5B with P-loop and autoimmune mutations
Schematic drawing of WRR5A and WRR5B proteins containing NB-ARC P-loop mutations and the *WRR5B* (*chs3-2D)* autoimmune mutation WRR5B-C1340Y. Domain predictions were performed using SMART and InterProScan (Jones, P. et al., 2014; Letunic, Ivica, Doerks and Bork, 2015). Annotations are from the start to the stop codon. Scale bar represents 100 amino acids.

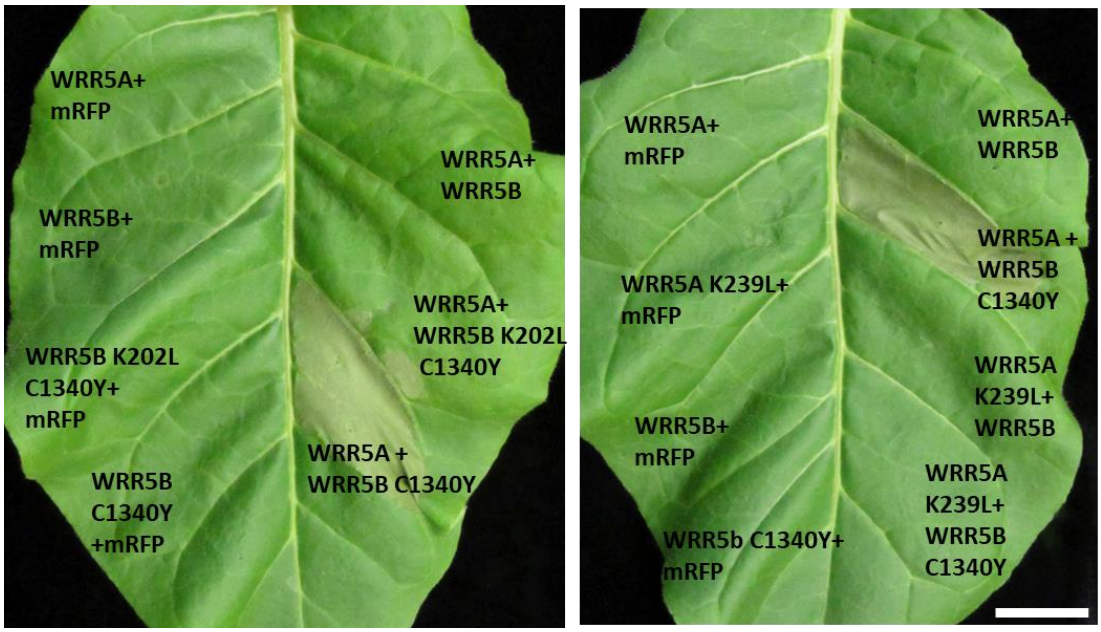


Figure 3.7: The P-loops of WRR5A and WRR5B are both required for WRR5A and WRR5B mediated immunity

Images of cell death assays following transient expression and agro-infiltration of *WRR5A*, *WRR5B* and the autoimmune allele encoding WRR5B-C1340Y in combination with P-loop mutants of WRR5A -K239L and WRR5B-K202L constructs in combination with each other and negative transgenic control construct monomeric Red Fluorescent Protein (mRFP) in segments of four week old *Nicotiana tabacum* leaves. Images were taken three days post infiltration with *Agrobacterium tumefaciens* strain GV3101. Scale bar represents 2cm.

DAR5 is the underlying gene for *WRR7* mediated resistance

The QTL analysis revealed three different candidate genes in the *WRR7* locus that could be responsible for AcEM2 resistance. All the candidates identified in this locus (*NRG1.1, NRG1.2* and *DAR5*) encode CC_R type resistance proteins and the other two resistance loci (*WRR4* and *WRR5*) encode TNL type of resistance proteins. TNLs activate disease resistance through lipase like protein EDS1 (Wagner et al., 2013). However, CNL or CC_R class of resistance genes operate independently of EDS1 (Wiermer, Feys and Parker, 2005). We therefore tested Col-*eds1.2* mutant plants with *A. candida* isolate AcEM2. We found Col-*eds1.2* plants were fully resistant to AcEM2 and showed similar resistance response as the RIL CW14 with functional $WRR7^{Col\text{-}0}$ only (Fig 3.8). We therefore conclude that *WRR7* mediated resistance is conferred by a CC_R type of resistance protein. Intriguingly, *DAR5 (At5g66630)* encodes a non-canonical resistance protein containing an RPW8 domain, partial NB-ARC, no LRR and an integrated LIM-Peptidase domain. Such an integrated domain was also found in WRR5B (Fig 3.6) making DAR5 a strong candidate for *WRR7*. To test this, we crossed Col-*eds1.2* with homozygous Col-*dar5.1* (SALK_068218C). We then identified homozygous Col-*eds1.2/dar5.1* (to be referred to as *wrr7*) line from F_2 individuals. We found this line to be fully susceptible to the pathogen (Fig 3.8) suggesting that *DAR5* is the underlying gene for *WRR7*. In addition, we transformed the AcEM2 susceptible *A. thaliana* ecotype Ws-2 as well as our Col-*eds1.2-wrr7* line with a *gWRR7* construct (Fig 3.9) to determine if either of their AcEM2 susceptibility phenotypes could be rescued.

We found that T_3 Ws-2 lines transformed with our genomic *WRR7* construct did not fully complement the resistant phenotype, but we did observe extensive cell death in these lines suggesting that *WRR7* activated a mild immune response that wasn't able to provide full resistance (Fig 3.10). In contrast, T_3 Col-*eds1.2-wrr7* mutant lines were complemented by *gWRR7,* with most lines showing a cell death response with no pustule development and some of the lines showing regions of cell death with reduced pustule development (Fig 3.10). Therefore, we have confirmed that *DAR5* is the causal gene for the *WRR7* phenotype and will from this point onwards refer to this gene as *WRR7*.

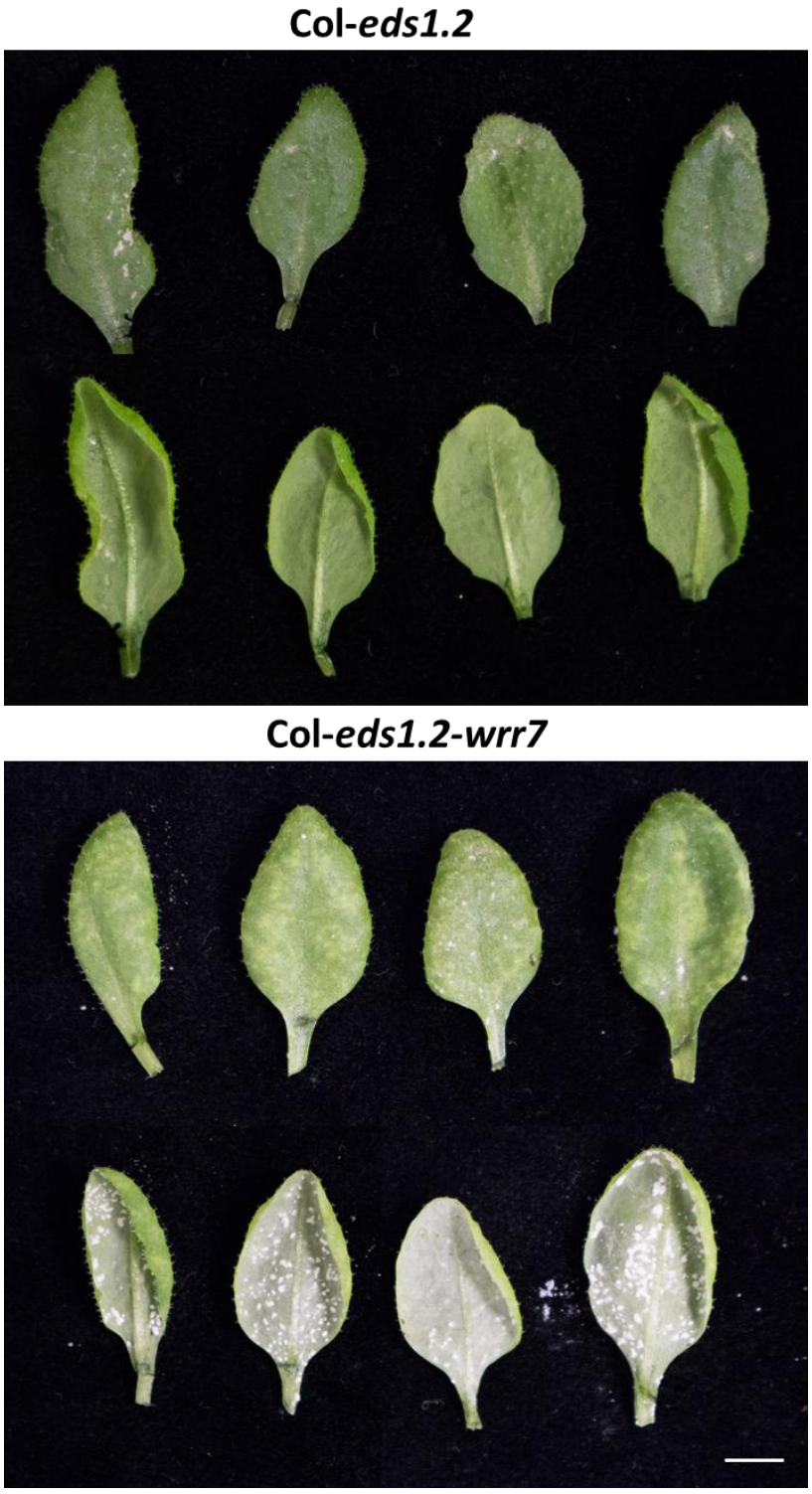


Figure 3.8: Col-*eds1.2-wrr7* lines are susceptible to AcEM2
Adult leaf phenotype images of Col-*eds1.2* and Col-*eds1.2-wrr7* lines 10 days post infection with *Albugo candida* isolate AcEM2. Scale bar represents 1cm.

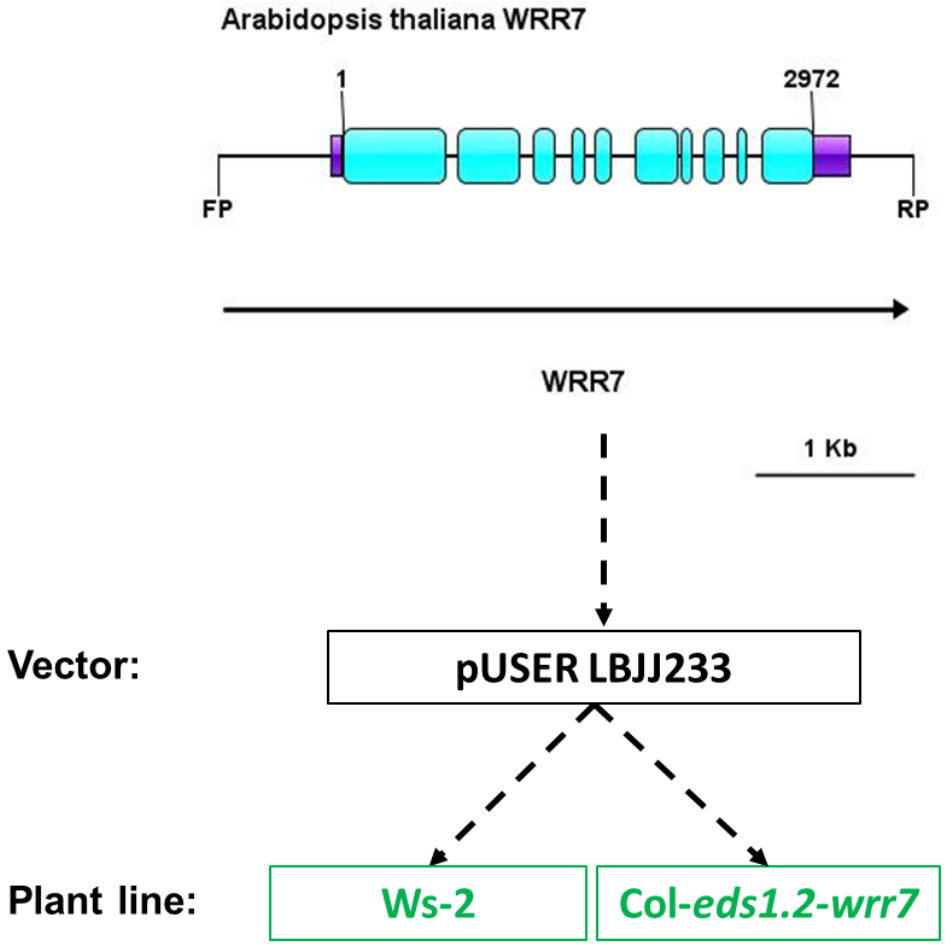


Figure 3.9: *WRR7* Cloning schematic
A - Diagram showing the cloning of *WRR7*$^{Col\text{-}0}$ (*At5g66630*) , blue rectangles depict the position of the exons, purple boxes show the untranslated regions, the gene is numbered from the start to the stop codon and the position of the primers in the promoter (1993 bp upstream of the start codon) and terminator (635 bp downstream of stop codon) regions are shown below the schematic diagram. Scale bar represents 1Kb. The *WRR7* genomic fragment was then cloned into the pUSER vector *LBJJ233* and then transformed into *Arabidopsis thaliana* lines Ws-2 and Col-*eds1.2-wrr7*.

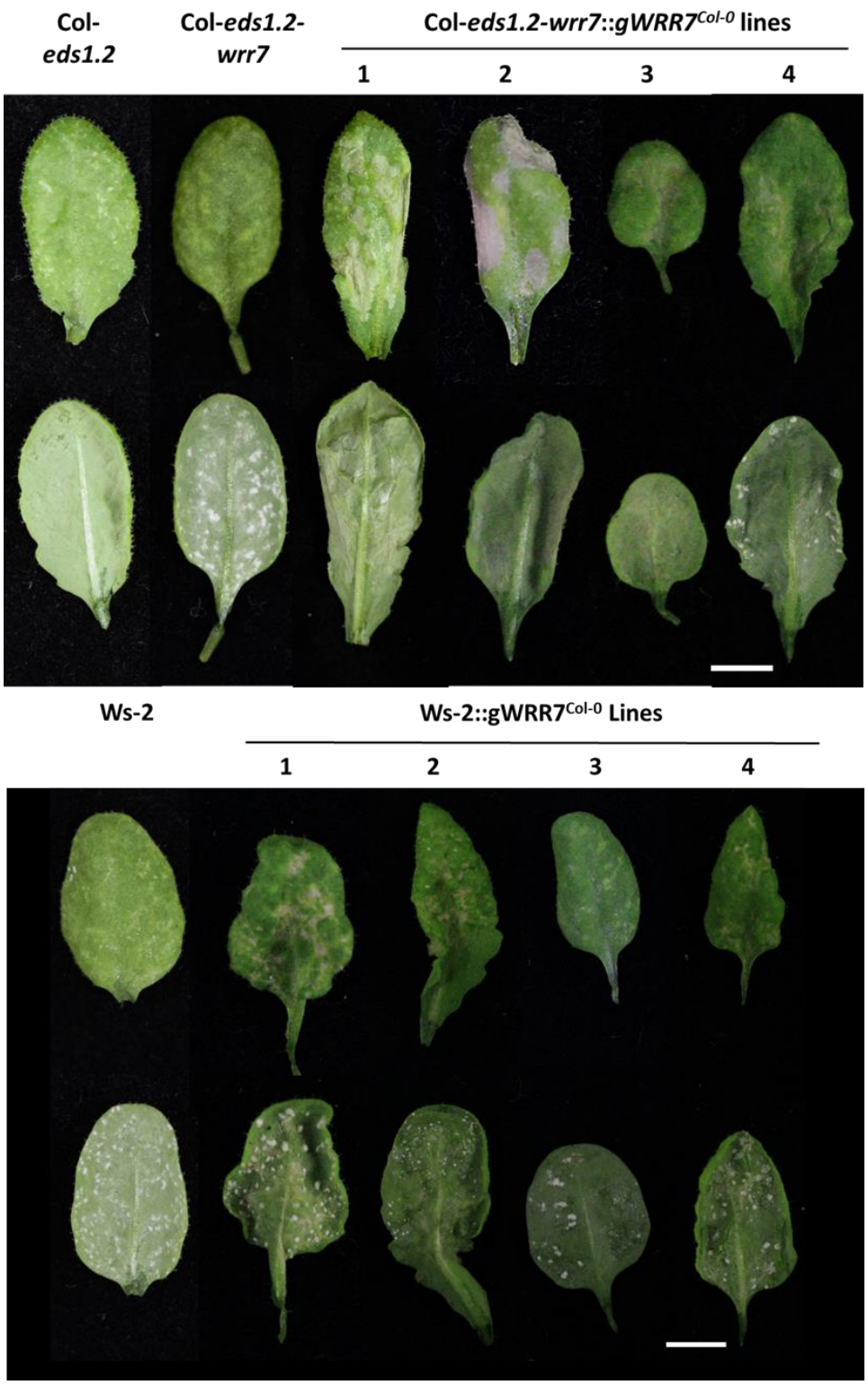


Figure 3.10: *WRR7* is able to complement the susceptible phenotype of Col-*eds1.2-wrr7* plants but not Ws-2 plants

Phenotypic images of *Arabidopsis thaliana* Col-*eds1.2-wrr7* and Ws-2 T_3 lines transformed with *gWRR7*. Leaf images were taken 10 days post infection (dpi) with *Albugo candida* isolate AcEM2, upper images show adaxial leaf surface and bottom panel shows the abaxial surface.

Discussion

Arabidopsis thaliana ecotypes display contrasting susceptibility phenotypes to the biotrophic Oomycete pathogen *Albugo candida*. The reference *A. thaliana* ecotype, Col-0, is fully resistant to *A. candida* race 4 isolate AcEM2, whereas the Ws-2 ecotype is fully susceptible. Recombinant inbred lines (RILs) generated from these two ecotypes were used to identify genes from the Col-0 genome that cause the resistance phenotype between the different ecotypes. Genotypic analysis of the RIL CW20 revealed that this line harbours only the previously reported TNL *WRR4A*, which is known to cause resistance to *A. candida* (Borhan, M. Hossein et al., 2008). In contrast genetic analysis of the RILs CW5 and CW14 showed that these lines harboured multiple resistance genes that could cause the resistance response to AcEM2. RIL CW5 contained *WRR5A* and *WRR5B*, two resistance genes that have previously been shown to both be required to cause an autoimmune response in transient expression analysis in *Nicotiana benthamiana* when the autoimmune allele of *WRR5B* (*chs3-2D*) is expressed (Xu et al., 2015). We also found that both of these two genes are required to cause the chlorotic resistance response, also observed in RIL CW5, to AcEM2. RIL CW14 harboured three CC_R type resistance genes that could have conferred the resistance response against AcEM2. Intriguingly, one of these resistance genes *DAR5* encoded a similar integrated LIM-Peptidase domain to *WRR5B*. We were subsequently able to show that the *WRR7* resistance phenotype was conferred by *DAR5* by mutational analysis and we renamed *DAR5* to *WRR7* based on its observed function. Therefore, in our analysis we have been able to identify that the Col-0 genome contains four resistance genes that are active against AcEM2. The function of three of these resistance genes *WRR5A, WRR5B* and *WRR7* are masked by the spatial epistasis provided by the much stronger resistance response conferred by *WRR4A*.

WRR5A and WRR5B are both required for the activation of autoimmunity imposed by the gain of function mutant *chs3-2D* (Bi, D. et al., 2011; Xu et al., 2015). The *chs3-2D* autoimmune phenotype is caused by a C to Y point mutation at amino acid site 1340, falling between the LIM and Peptidase domains of WRR5B (Bi, D. et al., 2011). Although WRR5A and WRR5B are known to operate as tandemly orientated, paired NLRs in activating autoimmunity in the *chs3-2D* mutant, their biological function was previously unknown. Our findings show that WRR5A and WRR5B

confer resistance against *A. candida* isolate AcEM2, in the *A. thaliana* Col-0 background. Paired NLR systems are known to cause resistance to several phytopathogens in both monocots and dicots, these include the *A. thaliana* RPS4/RRS1 system as well as *Oryza sativa* RGA4/RGA5 and Pik-1/Pik-2 paired NLR systems (Narusaka et al., 2009; Ortiz et al., 2017). In paired NLR systems, it is common for one of the NLRs to contain an extra integrated domain such as the WRKY domain of RRS1 or the heavy meatal associated domain RATX1 found in RGA5 and Pik-1. These domains act as the sensing modules of the NLR system and their role is to recognise the presence of pathogen effectors. The function of their paired NLR is to execute defence signalling following effector recognition. Therefore, the sensor-helper model was proposed suggesting that in paired NLRs the function of one of the NLRs it to detect the pathogens presence and the other enacts defence signalling (Cesari et al., 2014). The requirement of both WRR5A and WRR5B to stimulate an autoimmune response and the presence of an integrated domain in the architecture of WRR5B suggests that WRR5A and WRR5B work via the sensor-helper model, whereby WRR5B 'senses' an effectors presence through its integrated LIM-Peptidase domain and WRR5A triggers ETI (Eitas and Dangl, 2010; Wu, C.-H. et al., 2017). To activate defence the helper NLRs exchange ADP for ATP from the P-loop region of the NB-ARC domain (Leipe, Koonin and Aravind, 2004; Wang, J., Wang, et al., 2019). Therefore, the sensor NLR is postulated to not require an active P-loop region, a theory that was proven to be correct in the RPS4/RRS1 and RGA4/RGA5 systems where the P-loop of sensor NLRs RRS1 and RGA5 is not required for defence activation (Césari et al., 2014; Sohn et al., 2014). To further elucidate whether WRR5A and WRR5B operate by utilising a similar mode of action to the RPS4/RRS1 system, we generated P-loop mutants of both WRR5A and the autoimmune allele of WRR5B (WRR5B-*C1340Y*) and co-expressed these mutants in *Nicotiana tabacum*. In the RPS4/RRS1 paired NLR system, the executioner NLR RPS4 is able to stimulate an autoimmune response in the absence of the senor NLR RRS1 (Huh et al., 2017). In contrast in the WRR5A/WRR5B paired system, the executioner NLR WRR5A could not stimulate an autoimmune response in the absence of WRR5B. In other paired NLR systems such as the RPS4/RRS1 and RGA4/RGA5 paired NLR systems it has also been found that only the P-loop region of the executioner NLRs, RPS4 and RGA4 are required to stimulate immunity (Césari et al., 2014; Sohn et al., 2014). Again, the

WRR5A and WRR5B system differs in this respect as the P-loops of both WRR5A and WRR5B are required for defence activation. Therefore, we have identified that WRR5A and WRR5B activate defence using a novel mechanism that requires both NLRs to be in an 'activate' state to stimulate defence.

Interestingly, both WRR7 and WRR5B contain integrated LIM-PEP domains and are therefore part of the DA1 protein family. The DA1 protein family is defined as having the presence of a singular LIM domain associated with a zinc metallopeptidase domain (PF12315) and is found only in plants (Zhao, M. et al., 2014; Dong, H. et al., 2017). Currently, the majority of research carried out on DA1 family members has focused in their role in regulating seed and organ size (Li, Yunhai et al., 2008; Peng, Yuancheng et al., 2015; Wang, J.-L. et al., 2017). The presence of this integrated domain in two independent resistance genes conferring resistance to one isolate of *A. candida* strongly hints at the interaction of *A. candida* isolate AcEM2 effectors with the WRR5B and WRR7 LIM-peptidase domains. Integrated domains have been identified to be closely related to domains targeted by phytopathogen effectors and are therefore believed to act as integrated decoy domains (Sarris, P. F. et al., 2016). The presence of LIM-Peptidase in these two resistance genes indicates that *A. candida* isolate ACEM2 effectors are targeting the LIM-Peptidase domain of one or multiple non-resistance gene DA1 family members. Therefore, we speculate that *A. candida* targets and manipulates DA1 family members during its infection, potentially to manipulate the hosts cellular developmental processes to impose susceptibility by manipulating the growth: defence trade off that is observed in plants (Albrecht and Argueso, 2017). The integration of the LIM-Peptidase domain into the architecture of WRR5B and WRR7 therefore potentially mimics the pathogens effectors natural DA1 family member target and allows these resistance genes to sense the presence of *A. candida* via the interaction of an effector with the DA1 family.

In summary, the *A. thaliana* Col-0 genome harbours four resistance genes and three independent resistance mechanisms active against one race 4 isolate of *A. candida*. TNLs WRR5A and WRR5B which have previously been shown to cause autoimmunity when the autoimmune allele WRR5B-*C442Y* is transiently expressed in *N. tabacum,* also cause resistance to *A. candida* in *A. thaliana*. Furthermore, WRR7 a CC_R type of resistance protein also causes resistance against *A. candida*,

seemingly independent of a partner NLR. The presence of the same integrated LIM-Peptidase domain in both a TNL and CC_R protein and subsequent resistance against the same pathogen race suggests a common mechanism that facilitates the rapid evolution of new resistance genes.

Chapter 4

Mechanistic insights into *WRR5A* and *WRR5B* mediated immunity

Introduction

Plant resistance genes fall into two distinct classes, the TNLs which contain the TIR domains at their N-terminus and CNLs which do not contain a TIR domain and instead often have a coiled-coil domain (CC), and a sister-group of CNLs contain RPW8 domains at their N-terminus (Jones, J.D.G., Vance and Dangl, 2016). The downstream signalling mechanisms employed by each of these three groups is important for understanding how resistance genes activate plant immune responses. Of the three groups, we know more about how TNLs signal than CNLs and CC_Rs. TNLs are obligate on lipase like protein EDS1 to enact downstream signalling events that result in immune activation (Wiermer, Feys and Parker, 2005). Intriguingly, two groups of CC_Rs, the ADR1 and NRG families, have been found to act downstream of several TNLs via two distinct pathways requiring either PAD4 or SAG1010 (Castel et al., 2018; Qi et al., 2018; Lapin et al., 2019). TNLs acting through the EDS1-SAG101 complex require NRGs whereas TNLs acting through the EDS1-PAD4 complex signal through ADR proteins (Lapin et al., 2019). Two of the NLRs identified to cause resistance to *Albugo candida* isolate AcEM2 are WRR5A and WRR5B which are both TNLs. Therefore, it is likely that the resistance mechanism employed by WRR5A and WRR5B, requires NRGs to activate downstream immunity.

Tandemly orientated heterodimeric complexes of NLRs that are positioned in a head to head arrangement in the genome such as *WRR5A* and *WRR5B* are being found to operate as dual NLR systems able to recognise invading plant pathogens. One of the first such systems identified, was the *A. thaliana* resistance *to Pseudomonas syringae* 4 (RPS4) and resistance to *Ralstonia solanacearum* 1 (RRS1) paired NLR system (Narusaka et al., 2009). Dual NLR systems typically involve one canonical NLR and one Non-canonical NLR that contains an integrated domain, most commonly at its C-terminus, which mimics an authentic domain of a host protein that is targeted by a pathogen effector, such as the WRKY domain of RRS1 (Ma et al., 2018). Integrated

domains have been identified across multiple plant genomes (Kroj et al., 2016; Sarris, P. F. et al., 2016) and the integrated decoy model of pathogen recognition is increasingly being identified as a mechanism whereby plants can recognise a pathogens presence by incorporating a part of a protein targeted by a pathogen into an NLRs architecture (Kroj et al., 2016). The mode of action of paired NLR systems is only just beginning to be comprehended, initial studies into the action of paired NLRs revealed that the sensor NLR recognises the presence of a pathogen via direct interaction of an effector protein with the integrated decoy domain of the sensor NLR as with the RPS4-RRS1 system (Deslandes et al., 2003). However, more intricate indirect sensing mechanisms by dual NLR systems are being identified, for example Rice NLR Pii-2 contains an integrated Nitrate induced (NOI) domain which binds to the rice exocyst protein OsExo70-F3, a target of *Magnaporthe oryzae* effector AVR-Pii and can therefore indirectly sense the presence of the pathogen through the OsExo70-F3 interaction with AVR-Pii, following this recognition its partner NLR Pii-1 executes immune signalling (Fujisaki et al., 2017).

WRR5A and *WRR5B* are a pair of tandemly orientated TNLs also known as *CSA1* and *CHS3*. Previous studies have shown that WRR5B has a C1340 to Y1340 gain of function point mutation named *chs3-2D* which confers autoimmunity (Bi, D. et al., 2011) and that the autoimmunity stimulated by the *chs3-2d* mutant is obligate on partner TNL, WRR5A to elicit an autoimmune phenotype (Xu et al., 2015). We have shown that the biological function of WRR5A and WRR5B is to provide a resistance mechanism against phytopathogen *A. candida* (Fig 3.5). However, the exact mechanism that the WRR5 dual NLR system employs to detect the presence of *A. candida* is yet to be determined. In this chapter, we show that WRR5A and WRR5B form a heterodimeric complex in the cell prior to infection that is localised predominantly to the plasma membrane, following infection the executor protein WRR5A becomes enriched in the nucleus providing evidence that it is acting as the executor of immune functioning following *A. candida* infection.

The exact steps that NLRs employ to trigger resistance responses following activation are still poorly understood. NLRs are highly sensitive proteins that once activated can lead to detrimental developmental defects such as cell death and autoimmune responses due do the activation of effector triggered immunity (ETI). It is therefore highly important that NLRs are not misregulated and are maintained in

the correct concentrations and ratios within the cell. Many, dual NLR systems have therefore evolved to be located in the same loci and are often regulated by the same promoter region by being tandemly arranged in inverse orientation (Eitas and Dangl, 2010; Baggs, Dagdas and Krasileva, 2017; van Wersch and Li, 2019). This spatial localisation allows the co-regulation of paired NLRs but also can lead to the clustering on NLRs that can be epigenetically co-regulated leading to a co-ordinated defence response once activated (van Wersch and Li, 2019). The epigenetic co-regulation of NLRs during defence activation requires a mechanistic link between MTI and ETI. Here we identified chromatin interacting transcriptional regulators, which interact with the WRR5 complex and could be involved in transcriptional reprogramming of defence responses following infection by *A. candida*.

Results

Overexpression of *WRR5A* and *WRR5B* stimulates an immune response in *A. thaliana* Ws-2 after infection with AcEM2

WRR5A and WRR5B are both required to stimulate an immune response against *A. candida* in *A. thaliana* Ws-2 lines (Fig 3.5) and WRR5A is required for the WRR5B-*chs3-2d* mutant to stimulate an autoimmune response (Xu et al., 2015). It is also well known that TNL paired NLRs heterodimerise to form immune complexes that are crucial for the execution of their function (Huh et al., 2017). Therefore, we hypothesised that WRR5A and WRR5B proteins physically associate and that they operate as sensor- helper NLRs with WRR5B as the sensor with its integrated LIM-Peptidase domain and WRR5A as the executor of immune signaling.

To determine whether this was the case, we first overexpressed *WRR5A* and *WRR5B* using the 35S cauliflower mosaic virus promoter individually in *A. thaliana* ecotype Ws-2. To test whether, overexpression of both genes could stimulate an immune response when uncoupled, we individually cloned, and epitope tagged *WRR5A* and *WRR5B* with V5 (sequence from V protein of simian virus 5) and His-FLAG (HF) tags under the control of the 35S cauliflower mosaic virus promoter and octopine synthase terminator. Independent homozygous transgenic lines in the Ws-2 background overexpressing both *WRR5A* and *WRR5B* epitope tagged constructs were then challenged with AcEM2 to see whether they could stimulate the same immune response we had observed with Ws-2 plants transformed with a genomic construct containing both of these genes.

Overexpression of *WRR5A* in *A. thaliana* ecotype Ws-2 homozygous, independent transgenic lines, resulted in plants showing an autoimmune phenotype (Fig 4.1), this was unexpected as no autoimmune phenotype was observed by the overexpression of this gene in *N. tabacum* transient cell death assays (Fig 3.7). However, independent homozygous transgenic lines overexpressing *35S:WRR5B-HF* only did not show an autoimmune phenotype and were phenotypically indistinguishable from Ws-2 plants (Fig 4.1). Therefore, to generate double transgenic lines containing epitope tagged and overexpressed *WRR5A* and *WRR5B* we had to self *35S:WRR5B-HF* mono-transgenic lines to generate homozygous T_3 populations which were then transformed with *35S:WRR5A-V5* constructs as well as *35S:GUS-V5*. We found that

independent T_3 Ws-2 lines containing overexpressed *WRR5B-HF* and *GUS:V5* that did not contain *WRR5A* were unable to cause resistance to *A. candida* isolate AcEM2 (Fig 4.1). However, double transgenic lines overexpressing both *WRR5A-V5* and *WRR5B-HF* resulted in chlorotic resistance, similar to that observed in RIL CW5 and *gWRR5A* and *WRR5B* in Ws-2 lines (Fig 4.1). Overexpression of *WRR5A-V5* in lines overexpressing *WRR5B-HF* did not result in the autoimmune phenotype observed in single *WRR5A-V5* transgenic lines (Fig 4.1), suggesting that *WRR5B* is involved in the inhibition of immunity caused by *WRR5A* in *A. thaliana*. Therefore, we determined that overexpression of *WRR5A-V5* and *WRR5B-HF* in Ws-2 was able to activate an immune response that was similar to the response observed in the RIL containing just the WRR5 loci. The generation of similar immune phenotypes by lines overexpressing *WRR5A* and *WRR5B* to lines transformed with *gWRR5A* and *WRR5B* means that studying lines overexpressing these two proteins is functionally relevant.

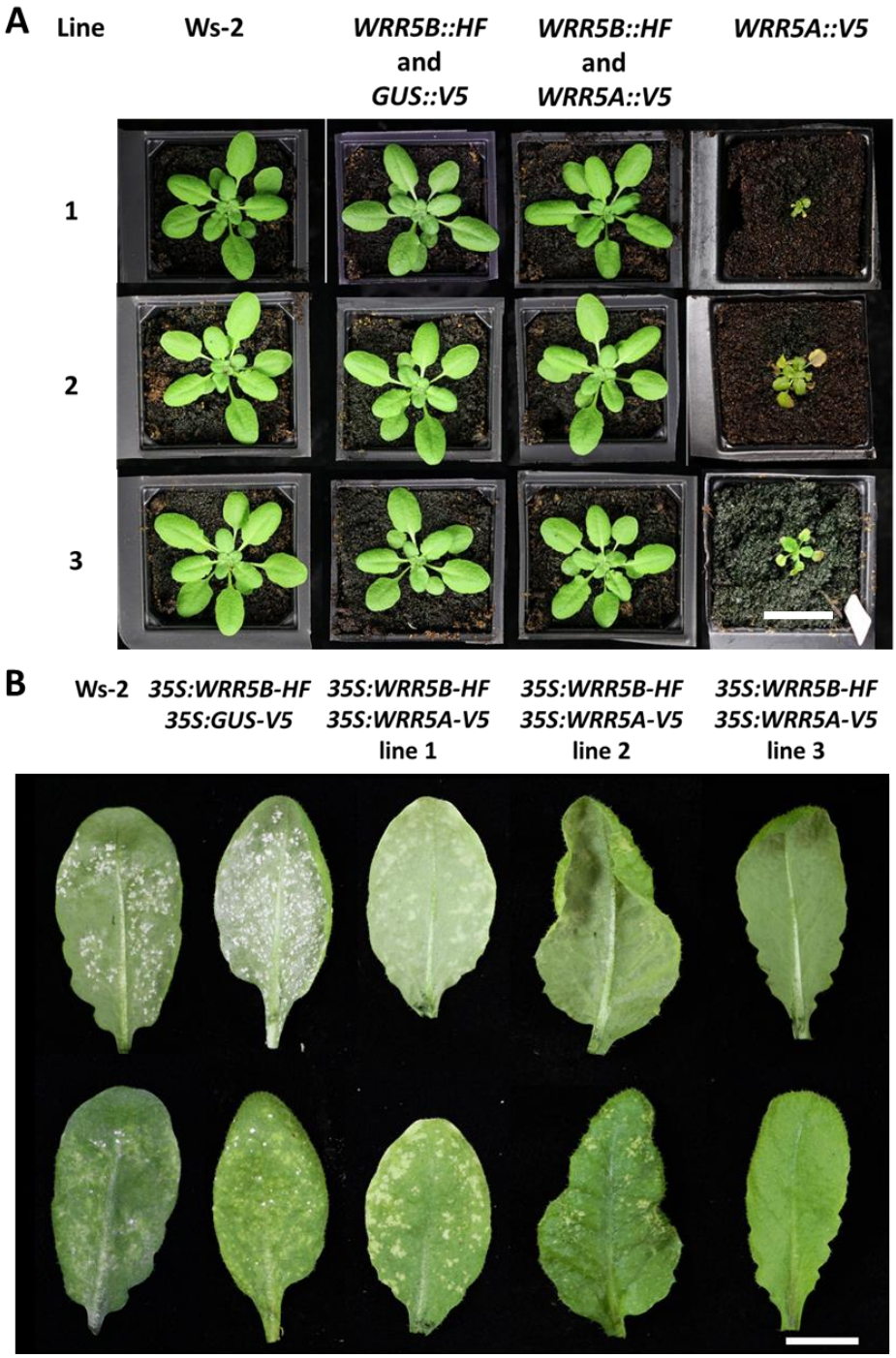


Figure 4.1: Overexpression of *WRR5A* and *WRR5B* stimulates an immune response in Ws-2 following *A. candida* infection

Phenotypic images of T_3 *A. thaliana* Ws-2 plants overexpressing epitope tagged *WRR5A*, *WRR5B* and *GUS* constructs driven by the 35S cauliflower mosaic virus promotor. A) Pre-infection phenotypes after 4 weeks of growth of three independent homozygous T_3 transgenic lines containing either *35S:WRR5A-V5*, *35S:WRR5B-HF*, *35S:GUS-V5* or combinations of these genes. Scale bar represents 2 cm. B) Post-infection phenotypes of three independent homozygous T_3 transgenic lines containing both *35S:WRR5A-V5* and *35S:WRR5B-HF* as well as a transgenic control line containing *35S:WRR5B-HF* and *35S:GUS-V5*. Images were taken of 4-week-old leaves 8 days post infection with *Albugo candida* isolate AcEM2. Scale bar represents 1 cm.

WRR5A and WRR5B form a heterodimer in *A. thaliana*

To test our hypothesis that WRR5A and WRR5B form a heterodimeric complex in *A. thaliana* and that this interaction is required to cause resistance against AcEM2, we carried out co-immunoprecipitation experiments with our Ws-2 double transgenic lines overexpressing *WRR5A* and *WRR5B* epitope tagged with either V5 or HIS-FLAG tags. WRR5A shares close homology with TNL RPS4 (At5g45250) in the *A. thaliana* genome (Cevik et al., 2019; Van de Weyer et al., 2019). To confirm this finding, we performed a BLASTn search against the *A. thaliana* TAIR 10 genome using the *WRR5A* genomic sequence and confirmed that *RPS4* is the most similar sequence. We confirmed that *RPS4* showed the closest homology to *WRR5A* with an E-value of 1e-141 (Fig 4.3).

To determine whether WRR5A and WRR5B interact in their natural host, *A. thaliana,* we extracted proteins from four independent T_3 lines overexpressing *WRR5B-HF*, with *WRR5A-V5* that had previously been shown to confer resistance to *A. candida* isolate AcEM2 (Fig 4.1) as well as two control lines overexpressing *WRR5B-HF* with *GUS-V5*. We then performed reciprocal co-immunoprecipitation experiments using both V5 and FLAG affinity bound antibody beads. Our co-immunoprecipitation experiments revealed that WRR5B-HF associated with WRR5A-V5 but not with GUS-V5 in *A. thaliana* in co-immunoprecipitation experiments using both V5 and FLAG antibody beads (Fig 4.2).

We then demonstrated that the interaction of WRR5A and WRR5B is highly specific by performing co-immunoprecipitation experiments with WRR5A-GFP, WRR5B-HF and RPS4-GFP which shows close homology with WRR5A. These proteins were transiently co-expressed in *N. benthamiana* and Co-immunoprecipitated using the GFP-Trap method. We found that WRR5B-HF could only associate with WRR5A-GFP constructs and not highly similar RPS4-GFP or GFP (Fig 4.3). Therefore, WRR5A and WRR5B association is highly specific.

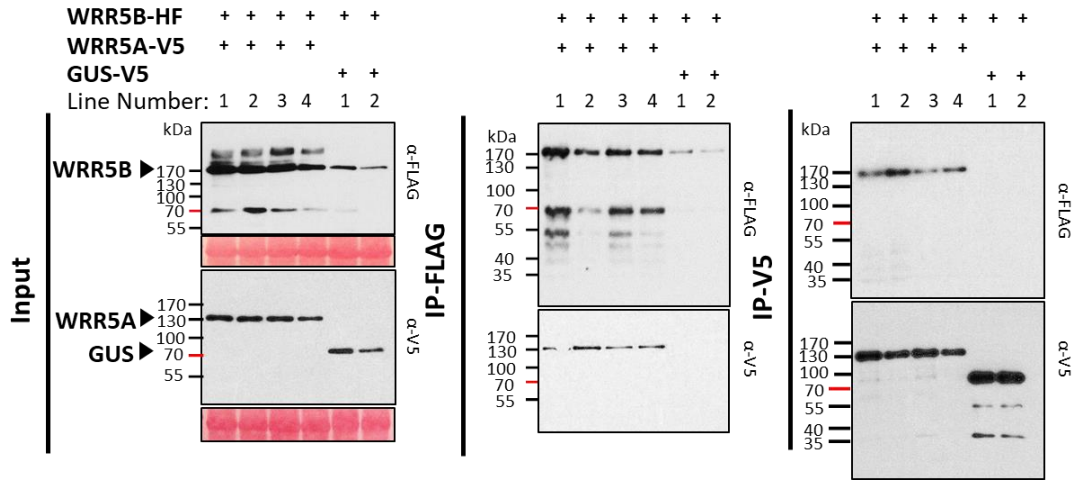


Figure 4.2: **WRR5A and WRR5B form a heterodimeric complex in *Arabidopsis thaliana***

Immunoblot results of total protein (input) and α-FLAG co-immunoprecipitated samples of constitutively overexpressed *WRR5B* tagged with His-FLAG tag and overexpressed *WRR5A* and *GUS* tagged with V5 in 4-week-old old *A. thaliana* double transgenic lines. Following visualisation, input sample membranes were stained with Ponceau-S and the RUBISCO band is shown below their respective blots. Expected proteins sizes are indicated with a triangle in the input blots.

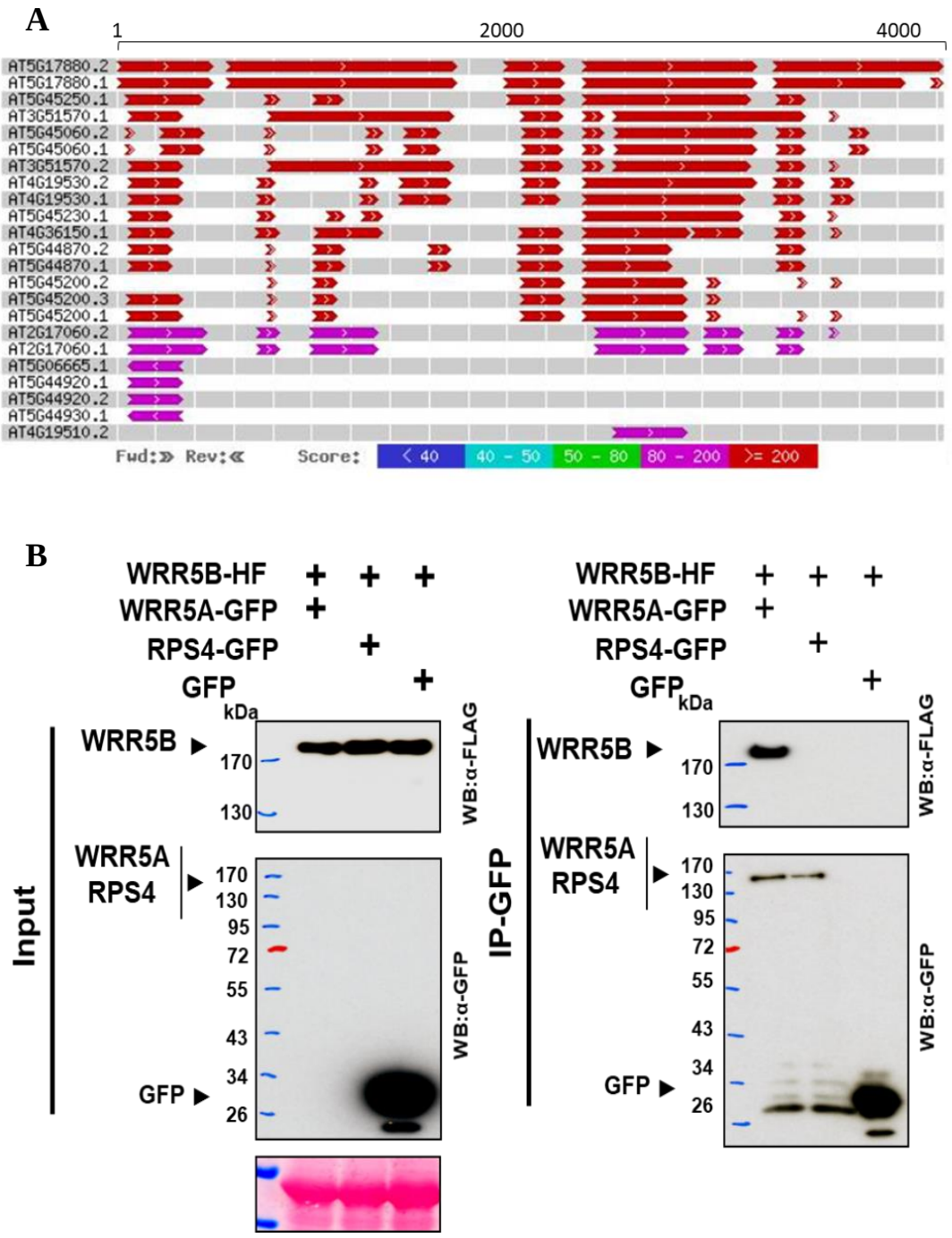


Figure 4.3: WRR5B associates with WRR5A-GFP but not RPS4-GFP in *Nicotiana benthamiana*

A) Alignment of *Arabidopsis thaliana* NLR nucleotide sequences showing closest homology (determined by BLASTn) with WRR5A (At5G17880). Scale shows nucleotide position.

B) Immunoblots of total protein (input) and αGFP co-immunoprecipitated proteins, from *N. benthamiana* leaves transiently overexpressing *WRR5B-HF*, *WRR5A-GFP*, *RPS4-GFP* and *GFP*, leaves were harvested 3 days post agro-infiltration. Following visualisation, input sample membranes were stained with Ponceau-S and the RUBISCO band is shown below their respective blots. Expected protein sizes are indicated with a triangle.

WRR5A and WRR5B localise to the plasma membrane

WRR5A and *WRR5B* are two resistance genes that are both required to cause resistance to *Albugo candida* race AcEM2 in *A. thaliana* ecotype Col-0. Both WRR5A and WRR5B proteins form a heterodimer *in planta* (Fig 4.2). However, it is unknown how they sense the presence of *A. candida* or how they activate disease following pathogen recognition. We believe that *WRR5A* and *WRR5B* operate by the sensor-helper model of NLR activation, whereby one NLR with an integrated domain (in this case *WRR5B*) recognises the pathogens presence and then activates or releases the helper NLR (*WRR5A*) which signals and activates defence responses. *WRR5A* alone causes an autoimmune response when overexpressed in *A. thaliana* Ws-2 lines, a phenotype that is rescued in the presence of *WRR5B* (Fig 4.1). Therefore, we speculate that WRR5B represses the immune activation mediated by WRR5A while they are in a heterodimeric complex and then releases WRR5A upon infection, resulting in immune signalling and the induction of defence related cell death. To test this hypothesis, we performed cellular fractionation experiments using *A. thaliana* Ws-2 lines constitutively overexpressing V5 and HF epitope tagged WRR5A and WRR5B in AcEM2 infected and mock infected plants. Leaves were harvested 3 days post inoculation with either water or a spore suspension of *A. candida* isolate AcEM2 and proteins were extracted using either nuclear or membrane fractionation. The obtained fractions contained proteins either enriched or depleted for nuclear proteins or fractions containing soluble or microsomal (membrane enriched) proteins. Cellular fractions were visualised using immunoblot analysis to identify their intracellular localisation in AcEM2 infected plants and in mock infected plants. In addition, we tagged both proteins with yellow fluorescent protein (YFP) and transiently expressed these proteins in *N. benthamiana* and visualised them using confocal microscopy to confirm the proteins localisation in the absence of a pathogen.

Our results revealed that both WRR5A-YFP and WRR5B-YFP are localised to the plasma membrane when transiently expressed in *N. benthamiana* (Fig 4.4), both proteins were visible in the plasma membrane, showing similar localisation to plasma membrane associated protein BAK1 and showed the presence of Hechtian strands after plasmolysis. We also found that WRR5A-YFP localises to the nucleus in some cells in transient expression experiments in *N. benthamiana*, however

WRR5B-YFP did not show any accumulation in the nucleus. To confirm that the localisation of WRR5A and WRR5B is consistent between *N. benthamiana* and *A. thaliana* we performed cellular fractionations of *A. thaliana* Ws-2 T_3 transgenic lines overexpressing *WRR5A-V5* and *WRR5B-HF*. Cellular fractions were confirmed by visualisation with native antibodies that bind to plasma membrane associated protein BAK1 and nuclear associated protein Histone H3. In contrast to transient expression analysis in *N. benthamiana*, cellular fractionation from *A. thaliana* revealed that both WRR5A-V5 and WRR5B-HF were present in both the nuclear enriched fraction as well as the microsomal fraction in mock inoculated plants (Fig 4.5). We were also able to show that, WRR5A-V5 but not WRR5B-HF showed an increase in the nuclear fraction proceeding infection by AcEM2, supporting the hypothesis that WRR5A-V5 is released from the heterodimer it forms with WRR5B following infection and relocates to the nucleus where it activates defence signalling. Our results therefore show that both WRR5A and WRR5B are localised to the plasma membrane and that WRR5A is also present in the nucleus and increases its nuclear presence following infection by AcEM2. Fluorescently labelled WRR5B does not accumulate in the nucleus when transiently expressed in N. benthamiana however WRR5B-HF was present in the nuclear fraction of our immunoblot analysis. Therefore, whether both WRR5A and WRR5B are localised to the nucleus or just WRR5A is unclear and further experimental evidence is needed.

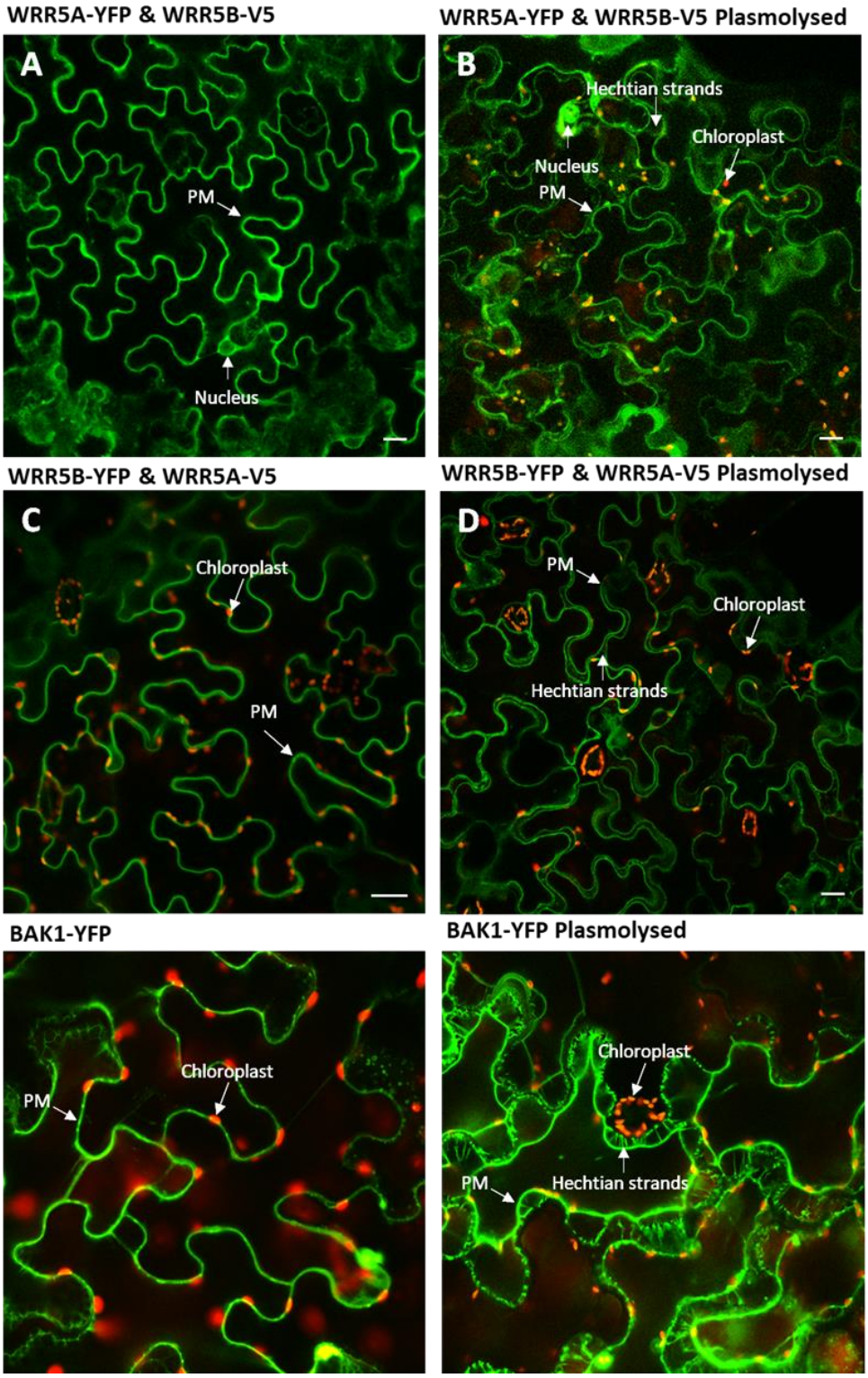


Figure 4.4: WRR5A-YFP and WRR5B-YFP localise to the Plasma membrane in *Nicotiana benthamiana*

Confocal microscopy images of transiently overexpressed YFP epitope tagged constructs in *Nicotiana benthamiana* leaf epidermal cells 2 days after infiltration with *Agrobacterium tumefaciens* strain GV3101. Red shows chloroplast autofluorescence and green shows YFP localisation. Scale bars represent 20µm. Examples of plasma membrane (PM), chloroplast, nucleus and Hechtian strands are labelled in each image.

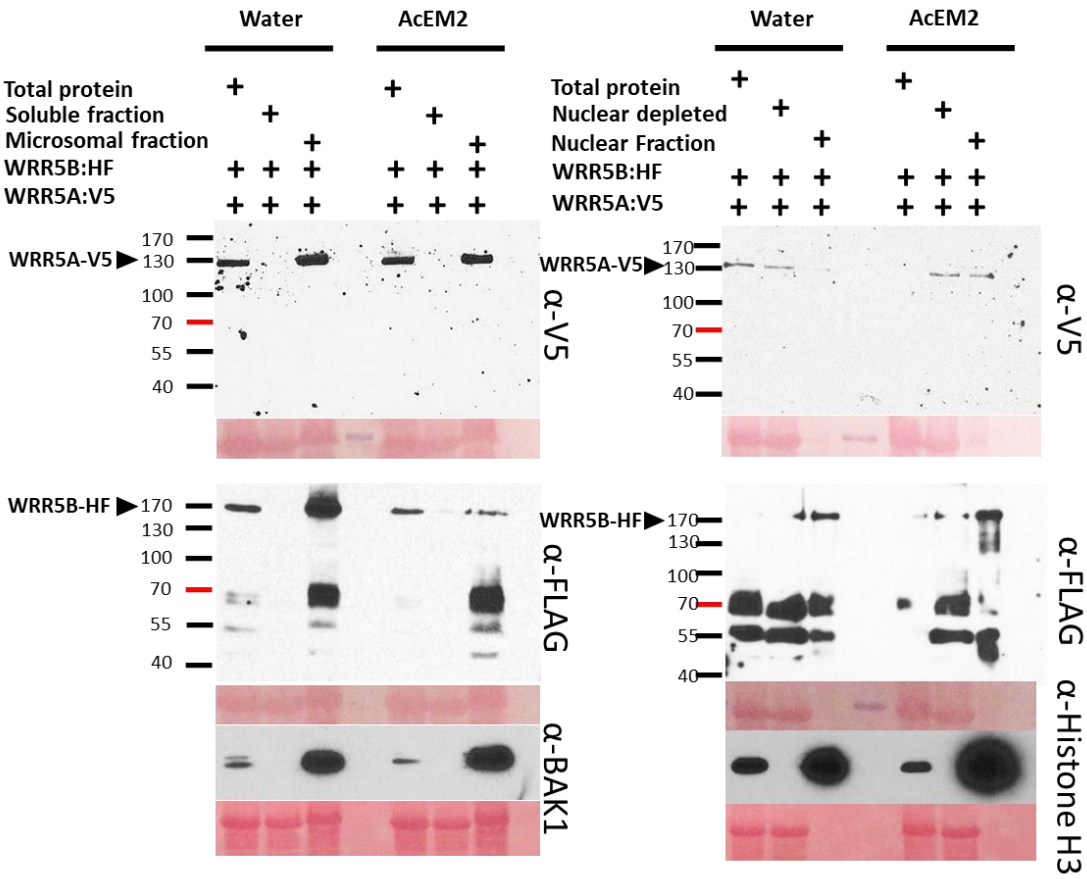


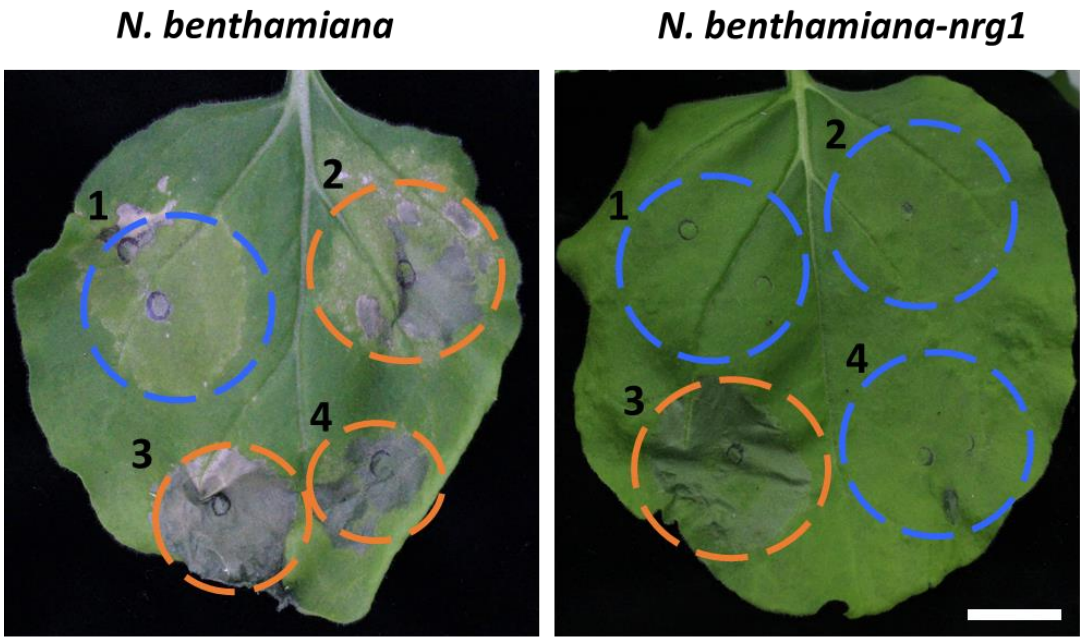
Figure 4.5: WRR5A and WRR5B localisation in *Arabidopsis thaliana*
Immunoblot results of proteins extracted from nuclear and microsomal fractions of double transgenic *A. thaliana* Ws-2 lines overexpressing *WRR5A-V5* and *WRR5B-HF*, 3 days post infection with water or with *A. candida* isolate AcEM2. Proteins were visualised using HRP-conjugated α-FLAG or α-V5 antibodies. Fractionated samples were visualised with α-BAK1, a plasma membrane associated marker and α-Histone H3 a nuclear marker. Following visualisation, membranes were stained with Ponceau-S and the RUBISCO band is shown below their respective blots. The expected WRR5A and WRR5B protein size is marked with a triangle.

WRR5A and *WRR5B* mediated immunity requires helper NLRs

NRG1 has been identified as a downstream signaling component of TNL resistance genes and some CNLs (Castel et al., 2018). *WRR5B* has a known gain of function autoimmune mutant *chs3-2D* caused by a G to A substitution resulting in a non-synonymous C1340Y mutation between the LIM- and Peptidase domains (Bi, D. et al., 2011). Therefore, to test whether *NRG1* was required for the autoimmune response elicited by *WRR5B-C1340Y* we obtained a *N. benthamiana-nrg*1 knock out line from (Castel et al., 2018) and transiently expressed both *WRR5A* and *WRR5B-C1340Y* in wild type *N. benthamiana* and *N. benthamiana-nrg1* to determine whether *NRG1* was required for the activation of an immune response mediated by the *WRR5A* and *WRR5B* complex. In this experiment we also tested if an autoimmune mutant of *WRR7* (Chapter 6) also requires a functional NRG1 in *N. benthamiana*. In addition to NRGs, ADR1 CC_Rs have also been shown to be involved in NLR downstream signalling (Lapin et al., 2019). To further test whether NRGs and ADR1 proteins are involved in WRR5 immune signalling in *A. thaliana*, we crossed the RIL CW5 (*wrr4/WRR5A&WRR5B/wrr7*) with triple knock out lines of the *nrg* (Wu, Z. et al., 2018) and *adr1* genes (Dong, O.X. et al., 2016) in the Col-0 background. F_2 seedlings from these crosses were then screened with *A. candida* isolate AcEM2. We would expect to identify AcEM2 susceptible seedlings at a ratio of 1:64 in the CW5 x Col-*nrg triple* mutant crosses as there are three segregating loci of interest (*WRR4*, *WRR7* and *NRG* loci) and at a ratio of 1:1024 for the CW5 x *Col-adr1 triple* which has five segregating loci as *ADR1, ADR1-L1* and *ADR1-L2* are present at separate loci.

We found that transiently expressed *gWRR5A* and *gWRR5B* did not stimulate an immune response in *N.benthamiana-nrg1* plants obtained from (Castel et al., 2018) or *N. benthamiana* wild type plants (Fig 4.6). In contrast, transiently expressed *WRR5B-C1340Y* co-expressed with *WRR5A* was able to induce a hypersensitive response in wild type *N. benthamiana* but not in *N. benthamiana-nrg1* (Fig 4.6). Whereas an autoimmune allele of WRR7 (chapter 6) was able to induce a hypersensitive response in both *N. benthamiana* and *N.benthamiana-nrg1* (Fig 4.6). We also confirmed that TNL *WRR4A* was unable to elicit an immune response in the presence of the recognised *A. candida* effector CCG28 in *N. benthamiana-nrg1* plants (Cevik et al, in prep). Furthermore, we identified seedlings that were partially

susceptible to AcEM2 in F_2 populations derived from the CW5 x Col-*adr1 triple* cross (Table 1 and Fig 4.7). However, no susceptible seedlings were identified in F_2 populations derived from the CW5 x Col-*nrg triple* cross (Table 1). Therefore, our data suggests that *NRG1* is required for *WRR5A* and *WRR5B* mediated autoimmune response in *N. benthamiana* but is not required to activate an immune response against *A. candida* in *A. thaliana* and intriguingly ADR1 proteins are required for WRR5 mediated immunity against *A.* candida in *A. thaliana*.



Infiltration number	Constructs infiltrated	Response in *N. benthamiana* (WT)	Response *in N. benthamiana-nrg1*
1	WRR5A and WRR5B		
2	WRR5A and WRR5B-C1340Y		
3	WRR7-C442Y		
4	WRR4A and CCG28		

Figure 4.6: WRR5A and WRR5B require NRG1 to activate defence responses
Cell death assays in *Nicotiana benthamiana* wilt type (WT) plants and *N. benthamiana-nrg1* plants agro-infiltrated *Agrobacterium tumefaciens* strain GV3101 transformed with overexpressed constructs outlined in the table. Cell death was observed in infiltrations shown in orange and no response was observed in infiltrations circled blue. A summary of the results is shown in the table. Scale bar represents 2cm.

Table 4.1: AcEM2 susceptible F_2 plants from CW5 x Col-*nrg* triple mutant and CW5 x Col-*adr1* triple mutant crosses are identified

Table of resistant: susceptible phenotypes observed in F_2 populations of crosses derived from WRR5 single loci containing recombinant inbred line CW5 (from original Col-5 x Ws-2 cross) crossed with nrg or adr1 triple mutant lines.

F2 line	Estimated number of seedlings screened	Number of susceptible seedlings identified	Glabrous segregating
CW5 x *nrg* triple	6000	0	Yes
CW5 x *adr1 L1/L2*	10,500	26	Yes

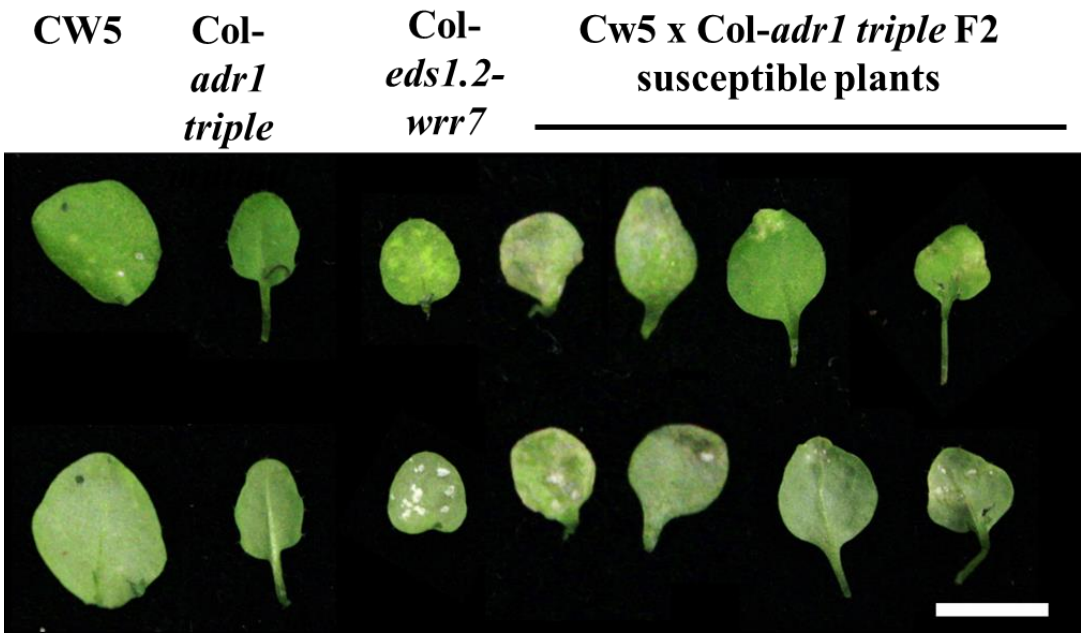


Figure 4.7: CW5 x *adr1 triple* mutant F2 plants show susceptible phenotype to AcEM2

Phenotypic images of *Arabidopsis thaliana* seedlings derived from a CW5 x Col-*adr1 triple* cross 10 days post infection with *Albugo candida* isolate AcEM2. Scale bars represents 0.5cm.

IP-MS identifies WRR5A and WRR5B interacting proteins

No proteins are currently known to interact with WRR5A or WRR5B. Therefore, to identify any potential interacting partners with the WRR5 complex or with WRR5A and WRR5B individually, we performed an immunoprecipitation-mass spectrometry analysis (IP-MS) using double transgenic *A. thaliana* Ws-2 lines overexpressing *WRR5A-V5* and *WRR5B-HF*. Three independent homozygous T_3 transgenic lines overexpressing *WRR5A-V5* and *WRR5B-HF* were grown and seedlings were collected and combined prior to protein extraction. Both HF and V5 beads were then exposed to the protein extract from these three lines, washed and analysed using liquid-chromatography mass spectrometry.

IP-MS identified 210 proteins following immunoprecipitation (IP) with V5 beads and 125 with FLAG beads incubated with protein extracts from *A. thaliana* lines overexpressing *WRR5A-V5* and *WRR5B-HF*. Of which, 66 proteins in the V5 IP and 45 in the FLAG IP were not detected in their respective V5 or FLAG bead controls treated with protein extract from wild type Ws-2 (Fig 4.8). Several abundant proteins were pulled down from the V5 and FLAG beads treated with non-transgenic control sample (60 and 51 respectively) and these proteins were used as a reference to remove any non-specific proteins that came down in the transgenic samples. In the transgenic samples we found that WRR5A and WRR5B both came down in the highest abundance in their reciprocal samples (Table 4.2 and 4.3), with 50 WRR5B-HF peptides in the WRR5A-V5 IP and 29 WRR5A-V5 peptides in the WRR5B-HF IP. Although, it has to be noted that seven WRR5B peptides were identified in the V5 IP control sample compared to 50 peptides in the WRR5A-V5 and WRR5B-HF sample but due to the large discrepancy in peptide number between the samples WRR5B was included in Table 4.2. WRR5A peptides were not present in the Ws-2 FLAG immunoprecipitation control samples, Therefore, our IP-MS data supports our earlier findings that WRR5A and WRR5B proteins form a heterodimeric complex in *A. thaliana* (Fig 4.2). Other than WRR5A and WRR5B, 14 other proteins were identified in the V5 IP (Table 4.2) and 11 in the FLAG IP (Table 4.3) that had a peptide count >1 and could be considered as interactors. In addition, 11 proteins (including WRR5A and WRR5B) were identified in both samples (Tables 4.4 and 4.5), these included proteins involved in protein folding such as Heat shock proteins (HSP), a lipoxygenase and a P-loop containing nucleoside triphosphate hydrolase. A

number of proteins were identified in one of either the WRR5A-V5 or the WRR5B-HF IP but not in both (Tables 4.2-4.5). These proteins could potentially interact with WRR5A or WRR5B individually. This group of proteins included a number of transcriptional regulators such as NOT3 and Chromatin remodelling complex (CHR2/BRM) which were present in the WRR5A-V5 IP but not in the WRR5B-HF IP (Tables 4.2 and 4.3). However, both these IP samples contained both WRR5A and WRR5B so we can not conclude that WRR5A alone interacts with the transcriptional regulators.

To identify whether the interactors identified by IP-MS had any common functions I performed a gene ontology (GO) term enrichment analysis with any proteins identified in either of the FLAG or V5 IPs. The GO term analysis showed that the proteins identified in the IP-MS experiments where implicated in responses to both abiotic and biotic stress (Fig 4.9). Therefore, the WRR5 complex is shown to be responsive to stimuli which corroborates our findings that it is a resistance gene active against *A. candida* but also previous findings that it is involved in the chilling response (Yang et al., 2010).

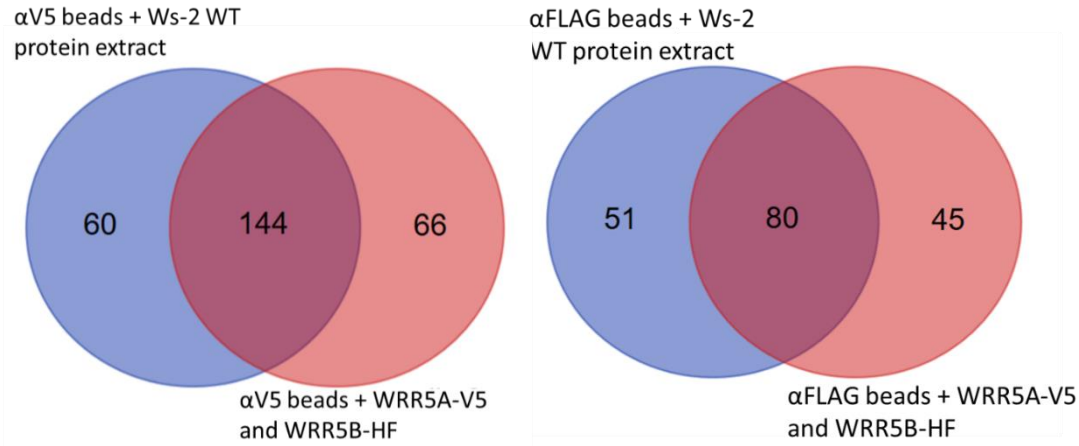


Figure 4.8: LC-MS results reveal proteins that are unique to WRR5A-V5 and WRR5B-HF IPs
Venn diagrams of LC-MS identified proteins immunoprecipitated from V5 and HF beads between control (Ws-2 protein extract) and protein extract from double transgenic *Arabidopsis thaliana* Ws-2 lines overexpressing both *WRR5A-V5* and *WRR5B-HF*.

Table 4.2: IP-MS V5 IPs identifies novel proteins associated with WRR5A
List of IP-MS identified proteins in *Arabidopsis thaliana* double transgenic lines overexpressing both *WRR5A-V5* and *WRR5B-HF*. Proteins were immunoprecipitated using V5 beads and proteins shown contained >1 peptide using a false discovery rate of 1% and were present in > 5 times the amount in the sample compared to the negative control. * Demarks proteins that were present in the control as well as the sample.

Accession	ATG number	Gene name	Score	Coverage	Proteins	Unique Peptides	Peptides	Area
A0A1P8BFL2	At5g17880	WRR5A/CSA1	816.36	64.05	3	74	74	1210437290
Q9FKN7*	At5g17890	WRR5B/CHS3	271.30	36.83	5	50	50	282578725
P22953	At5g02500	HSPC70-1 (heat shock cognate protein 70)	54.32	26.57	9	3	12	104713175.7
O65719	At3g09440	HSP 70 (heat shock protein 70)	41.06	20.80	5	1	10	103900924.6
P38418	At3g45140	LOX2 (Lipoxygenase 2)	33.45	11.72	2	8	8	20059702.79
Q93VB0	At2g40660	Nucleic acid-binding, OB-fold-like protein	12.47	18.25	1	4	4	21268754.15
Q94AW8	At3g44110	ATJ (DNAJ chaperone homologue 3)	6.42	11.90	3	4	4	9121070.583
Q9S9N1	At1g16030	HSP 70B	16.48	8.98	5	1	4	103372032.7
F4JL78	At4g28300	Formin-like protein	7.22	7.76	2	3	3	11578304.05
F4JWJ6	At5g18230	NOT3 (transcriptional regulator)	27.57	4.86	4	3	3	5080753.719
Q9SIH0	At2g36160	40S ribosomal protein S14-1	9.67	23.33	3	3	3	32840606.93
A8MRV1	At1g07660	Histone H4	2.46	23.26	2	2	2	7466146.688
P51818	At5g56010	HSP90-3	2.19	3.58	4	2	2	5160879.813
P56801	AtCg00770	30S ribosomal protein S8	6.87	14.93	1	2	2	22350260.28
Q6EVK6	At2g46020	BRM/CHR2 (Chromatin remodeling complex 2)	5.95	1.14	1	2	2	5847726.438
Q9LNU1	At1g20160	SBT5.2 (subtilisin-like serine protease)	7.53	3.38	2	2	2	6922115.336

Table 4.3: IP-MS HF IPs identifies novel proteins associated with WRR5B
List of IP-MS identified proteins in *Arabidopsis thaliana* double transgenic lines overexpressing both *WRR5A-V5* and *WRR5B-HF*. Proteins were immunoprecipitated using FLAG beads and proteins shown contained >1 peptide using a false discovery rate of 1% and were present in > 5 times the amount in the sample compared to the negative control.

Accession	ATG number	Gene name	Score	Coverage	Proteins	Unique Peptides	Peptides	Area
Q9FKN7	At5g17890	WRR5B/CHS3	294.19	36.21	9	49	49	916061134.4
A0A1P8BFL2	At5g17880	WRR5A/CSA1	105.66	27.85	3	29	29	220196976
P10896	At2g39730	RCA (Rubisco activase)	168.93	50	2	17	17	1196446055
O65719	At3g09440	Hsp 70 (heat shock protein 70)	22.76	14.02	5	1	8	129566890.4
B9DGT7	At1g50010	TUA2 (Tubulin alpha-2 chain)	11.29	10.44	3	1	4	62826085.75
P25858	At3g04120	GAPC (Glyceraldehyde-3-phosphate dehydrogenase C subunit 1)	13.41	14.79	3	4	4	104871135.9
P29511	At4g14960	TUA6	6.80	10.44	4	1	4	60632195.33
A0A2P2CLF9	AtMg01190	ATP1 (ATP synthase subunit 1)	5.75	4.93	3	1	3	136375285.2
P38418	At3g45140	LOX2 (Lipoxygenase 2)	4.12	3.79	2	3	3	14330582.02
P31167	At3g08580	AAC1 (ADP,ATP carrier protein 1)	9.00	6.04	3	2	2	11680223.38
Q94AW8	At3g44110	ATJ (DNAJ chaperone homologue 3)	0.00	6.19	3	2	2	7167311.953
Q9CAX6	At3g11510	40S ribosomal protein S14	4.53	16	3	2	2	49705405.53
Q9SYW8	At3g61470	LHCA2 (Photosystem 1 light harvesting complex gene 2)	14.50	12.06	1	2	2	73737755

Table 4.4: IP-MS HF IPs identifies novel proteins associated with both WRR5A and WRR5B

List of IP-MS identified proteins identified in *Arabidopsis thaliana* double transgenic lines overexpressing both *WRR5A-V5* and *WRR5B-HF*, proteins were immunoprecipitated using FLAG bead IP experiments and were also present in V5 bead IPs using a false discovery rate of 1%.

Accession	ATG number	Gene name	Score	Coverage	Proteins	Unique Peptides	Peptides	Area
Q9FKN7	At5g17890	WRR5B/CHS3	294.19	36.21	9	49	49	916061134
A0A1P8BFL2	At5g17880	WRR5A/CSA1	105.66	27.85	3	29	29	220196976
P38418	At3g45140	LOX2 (Lipoxygenase 2)	4.12	3.79	2	3	3	14330582
Q94AW8	At3g44110	ATJ (DNAJ chaperone homologue 3)	0.00	6.19	3	2	2	7167312
O65719	At3g09440	Hsp 70 (heat shock protein 70)	22.76	14.02	5	1	8	129566890
A0A2P2CLF9	AtMg01190	ATP1 (ATP synthase subunit 1)	5.75	4.93	3	1	3	136375285
A0A1P8ATD8	At1g73110	P-loop containing nucleoside triphosphate hydrolase	4.77	3.28	2	1	1	9452237
P43286	At3g53420	PIP2 (plasma membrane intrinsic protein 2)	0.00	3.14	2	1	1	30465989
Q6XJG8	At4g28470	26S Proteosome regulatory subunit S2 1B	0.00	1.68	2	1	1	14586463
Q94CE4	At1g70410	BCA4 (Beta carbonic anhydrase 4)	2.43	4.29	1	1	1	8990537
Q9M356	At3g61820	Aspartyl protease	1.78	1.86	1	1	1	15695825

Table 4.5: IP-MS V5 IPs identifies novel proteins associated with both WRR5A and WRR5B

List of IP-MS identified proteins identified in *Arabidopsis thaliana* double transgenic lines overexpressing both *WRR5A-V5* and *WRR5B-HF*, proteins were immunoprecipitated using V5 bead IP experiments and were also present in FLAG IPs using a false discovery rate of 1%. * Demarks proteins that were also present in the negative control in an abundance <5x the amount than in the sample.

Accession	ATG number	Gene name	Score	Coverage	Proteins	Unique Peptides	Peptides	Area
A0A1P8BFL2	At5g17880	WRR5A/CSA1	816.36	64.05	3	74	74	1210437290
Q9FKN7*	At5g17890	WRR5B/CHS3	271.30	36.83	5	50	50	282578725
O65719	At3g09440	Hsp 70 (Heat shock protein 70)	41.06	20.8	5	1	10	103900925
P38418	At3g45140	LOX2 (Lipoxygenase 2)	33.45	11.72	2	8	8	20059703
Q94AW8	At3g44110	ATJ (DNAJ chaperone homologue 3)	6.42	11.9	3	4	4	9121071
A0A2P2CLF9	AtMg01190	ATP1 (ATP synthase subunit 1)	2.14	2.17	3	1	1	25379539
A0A1P8ATD8	At1g73110	P-loop containing nucleoside triphosphate hydrolase	2.68	3.28	2	1	1	5306459
P43286	At3g53420	PIP2 (plasma membrane intrinsic protein 2)	1.61	3.14	2	1	1	25331403
Q6XJG8	At4g28470	26S Proteosome regulatory subunit S2 1B	2.29	1.12	2	1	1	4772666
Q94CE4	At1g70410	BCA4 (Beta carbonic anhydrase 4)	2.40	4.29	1	1	1	4747993
Q9M356	At3g61820	Aspartyl protease	1.66	2.07	1	1	1	11854176

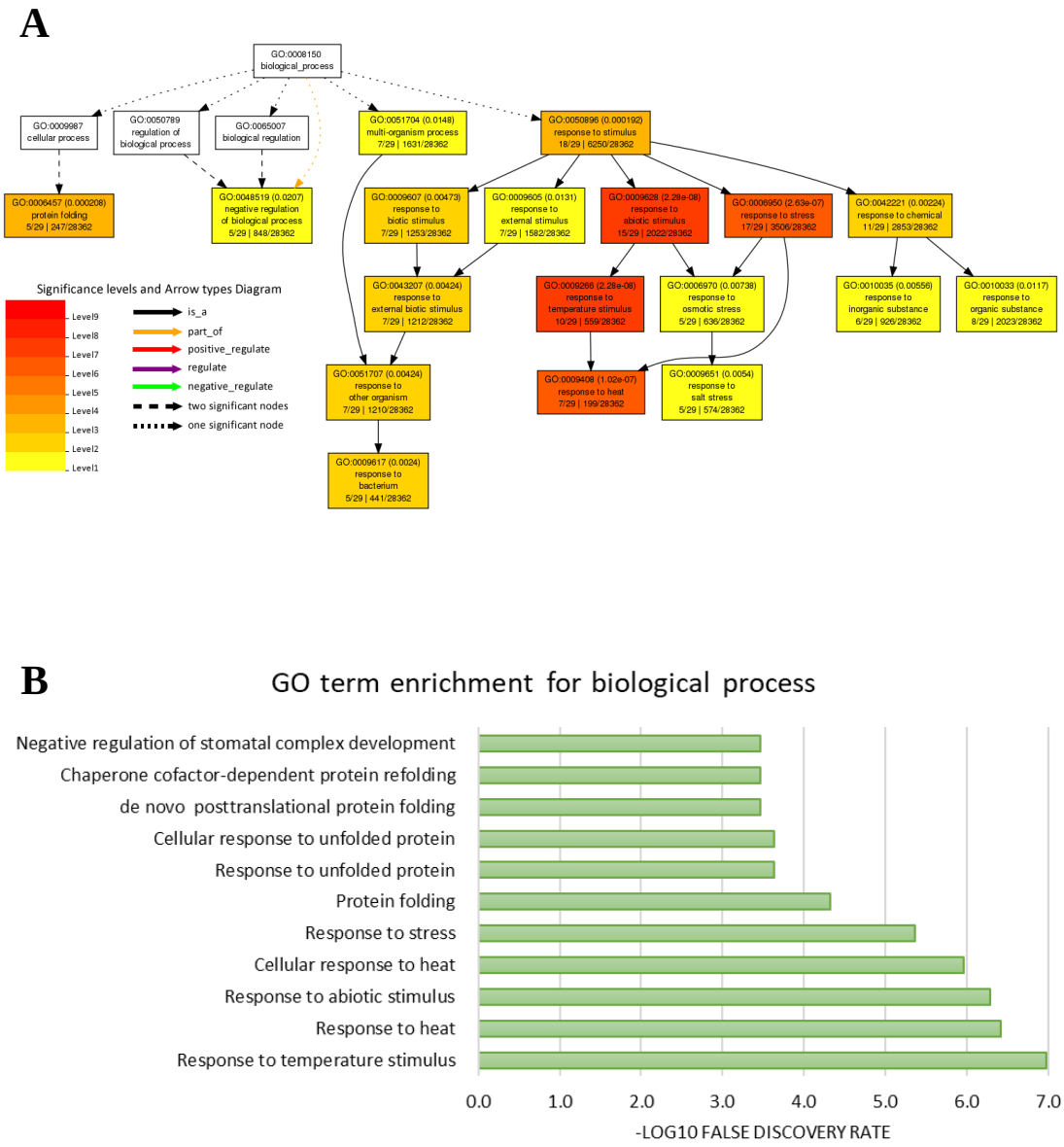


Figure 4.9: Gene ontology analysis of proteins associated with WRR5A and WRR5B by IP-MS

Gene ontology (GO) enrichment analysis of proteins identified with a peptide count >1 from V5 and HF IP-MS samples containing protein extract from *Arabidopsis thaliana* ecotype Ws-2 overexpressing *WRR5A-V5* and *WRR5B-HF* as well as proteins with a combined peptide count >1 that were identified in both samples. GO term analysis was performed using singular enrichment analysis and using agriGO software (Tian, T. et al., 2017) and compared to the latest *A. thaliana* database in TAIR 10 (Lamesch et al., 2011). GO terms were significant at a P-value of <0.05 threshold. A) GO term network of biological functions identified and B) Bar chart of the top 10 most enriched biological functions from the GO term analysis with the lowest false discovery rates.

Discussion

The two *A. thaliana* TNLs WRR5A and WRR5B form a paired defence system which has previously been shown to induce an autoimmune response if a particular site (C1340Y) in WRR5B was mutated (Bi, D. et al., 2011; Xu et al., 2015). However, other than the finding that this mutation could cause autoimmunity their biological function in plant defence was unknown. In the previous chapter, I demonstrated that *WRR5A* and *WRR5B* play an active role in the defence of *A. thaliana* accession Col-0 against the AcEM2 isolate of oomycete phytopathogen *A. candida* (Fig 3.5). In this chapter, I showed that WRR5A and WRR5B proteins form a heterodimer in *A. thaliana* and that both *WRR5A* and *WRR5B* are required for the immune response in *A. thaliana* against *A. candida* isolate AcEM2. In addition, we were able to further elucidate the signalling mechanism that the WRR5A and WRR5B immune complex employs to cause resistance to *A. candida* isolate AcEM2.

I employed several methods to determine how WRR5A and WRR5B function following an invasion by *A. candida*. Firstly, we wanted to determine the sub-cellular localisation of WRR5A and WRR5B before infection and then determine whether their localisation changed following infection by *A. candida*. To do this, we cloned and tagged both genes with C-terminal epitope tags and analysed their localisation when transiently expressed in *N. benthamiana.* We also constitutively expressed both genes in *A. thaliana* and used cellular fractionation experiments to determine their locations in the presence or absence of *A. candida*. The location of WRR5A and WRR5B was then determined using confocal microscopy and cellular fractionation and immunoblotting (Fig 4.4 and 4.5). We found that both proteins were present in the microsomal fraction but not the soluble fraction of uninfected *A. thaliana* protein samples (Fig 4.5). In addition, we showed by confocal microscopy that both proteins tagged with YFP were localised to the plasma membrane when transiently expressed in *N. benthamiana* and that WRR5A but not WRR5B also accumulated in the nucleus (Fig 4.4). Intriguingly, immunoblot analysis of cellular fractions revealed that WRR5A and WRRB were present in the nuclear enriched fraction but WRR5A shows an increase in this fraction after infection with AcEM2 (Fig 4.5). Therefore, we speculate that WRR5A dissociates from the plasma membrane and relocates to the nucleus following infection by *A. candida*. This mechanisms of activation is different to the RPS4 and RRS1 paired NLR system that

is predominantly localised to the nucleus (Huh et al., 2017). Therefore, I propose a system whereby WRR5A, potentially in complex with WRR5B relocates to the nucleus following infection, whether this translocation is undergone by both proteins or just WRR5A still needs further experimental validation.

We also found that WRR5A and WRR5B require the helper NLR NRG1 in *N. benthamiana* transient expression experiments in accordance with previous findings that NRGs are required for the immune function of TNLs (Castel et al., 2018). However, crosses between the RIL CW5 (which contains only the WRR5 resistance loci and not the WRR4 and WRR7 loci) and a triple T-DNA knock out line of the three NRG genes yielded no susceptible F_2 individuals suggesting that although NRG1 may be required for an autoimmune response the NRGs may not be required for the immune response elicited by AcEM2, exactly how these responses differ needs further experimentation. Another finding was that in *A. thaliana* the ADR1 proteins are required for the WRR5A and WRR5B mediated immune response to *A. candida* as susceptible F_2 individuals derived from a cross of CW5 x *adr1* triple mutant lines were identified again this finding was unexpected as the loss of ADR1 proteins has previously shown to have no impact on the autoimmune response conferred by the WRR5A and WRR5B complex (Dong, O.X. et al., 2016). Taken together these results suggest that the autoimmune response triggered by the WRR5B-*chs3*-2D mutation may operate via a different mechanism to the immune response triggered following *A. candida* infection, exactly how these two responses differ requires further experimental work.

Once the localisation of WRR5A and WRR5B was determined we wanted to identify other proteins that associate with the WRR5A and WRR5B complex that could shed light on their mechanistic operation. We employed an immunoprecipitation-mass spectrometry (IP-MS) analysis to determine any proteins that were associated with the WRR5 immune complex in *A. thaliana*. We identified 29 interacting proteins with the WRR5A and WRR5B complex that had a peptide count >1, from individual FLAG or V5 co-immunoprecipitated samples containing WRR5A-V5 and WRR5B-HF or that were present in both IP samples (Table 4.2-4.5). These included several proteins that regulate the transport and correct folding of proteins such as heat shock proteins and chaperones as well as plasma membrane channels and nuclear proteins that are involved in transcriptional reprogramming. One of the most interesting

proteins that came down in both the WRR5A-V5 and WRR5B-HF immunoprecipitated samples is the aquaporin Phosphatidylinositol (4,5)-bisphosphate (PIP2), a plasma membrane intrinsic protein that is used as a plasma membrane marker (Czech, 2000). PIP2 is also an important component of lipid signalling events and has been implicated in the regulation of stomatal guard cell closure and has been shown to be responsive to changes in Ca^{2+} homeostasis (Tuteja and Sopory, 2008; Byrt et al., 2017). Therefore, the co-immunoprecipitation of this protein backs up our finding that the WRR5A and WRR5B complex is plasma membrane localised but also may offer an interesting node at the plasma membrane that could be guarded by the WRR5A and WRR5B complex, allowing it to sense the presence of *A. candida*. The protein that was identified in the greatest abundance in both FLAG and V5 IPs was Lipoxygenase 2 (LOX2) with 11 peptides identified between the two samples (Table 4.4 and 4.5). Lipoxygenases are involved in the jasmonic acid (JA) phytohormone biosynthesis pathway and LOX2 is specifically involved in the production of Ca^{2+} responsive green leaf volatiles that are released following tissue disturbance (Mochizuki and Matsui, 2018). The presence of this protein in the WRR5A and WRR5B IPs was unexpected as it is predominantly localised to the chloroplast. However, LOX proteins are known to relocate to membranes following increases in Ca^{2+} concentration (Walther, Wiesner and Kuhn, 2004; Järving et al., 2012) and this mechanism of activation has already been postulated for AtLOX2 (Mochizuki and Matsui, 2018). Therefore, this represents another mechanism whereby the WRR5A and WRR5B complex could sense the presence of *A. candida* through its interaction with LOX2. This mechanism would operate by *A. candida* stimulating an increase in Ca^{2+} during invasion causing a Ca^{2+} induced relocation of LOX2 to the plasma membrane where it interacts with WRR5A and WRR5B complex stimulating an immune response (Fig 4.10).

Interestingly, we also identified two transcriptional regulators, NOT3 and Chromatin remodelling complex 2 (CHR2 also known as BRM) that were identified solely in the V5 bead IP but not in the FLAG IP (Table 4.2 and 4.3). Therefore, these proteins had a higher affinity to WRR5A than WRR5B. NOT3 is part of the Carbon catabolite repressor 4 CCR4-NOT complex which regulates transcription by altering chromatin structure (Lau et al., 2009; Arae et al., 2019). Chromatin remodelling complexes (CHR) are localised to the nucleus and epigenetically regulate genes by

modifying histones. CHR2/BRM is one of four Switch/Sucrose non-fermenting chromatin remodelling complexes in *A. thaliana* (Thouly et al., 2020). It has been shown to associate with the promoter regions of over 5000 genes, that are associated with responses to various stimuli, among the genetic regions associated with CHR2 is a region associated with WRR5A (Li, C. et al., 2016). CHR2 is known to associate with a plant specific H3K27me3 demethylase Relative of Early Flowering 6 (REF6) and together they act antagonistically to Polycombe group protein function by removing H3K27me3 marks (Li, C. et al., 2016; Thouly et al., 2020). The interactions of the WRR5 complex with transcriptional regulators provides an interesting mechanism that the WRR5 complex could utilise to stimulate defence responses through epigenetic reprogramming.

Our experiments have determined that WRR5A and WRR5B are present in the plasma membrane of *A. thaliana* prior to infection and that WRR5B plays an inhibitory role in suppressing the autoimmune response induced by WRR5A in the absence of a pathogen, by sequestering WRR5A in a heterodimeric complex. Following infection by *A. candida*, this complex then senses the presence of an as yet unidentified effector possibly through its association with PIP2 or LOX2 proteins in the plasma membrane, resulting in the dissociation of WRR5A from WRR5B. WRR5A, and potentially WRR5B then translocates to the nucleus where it interacts with transcriptional regulators such as CHR2 or NOT3 causing transcriptional reprogramming leading to a hypersensitive response of the infected cells.

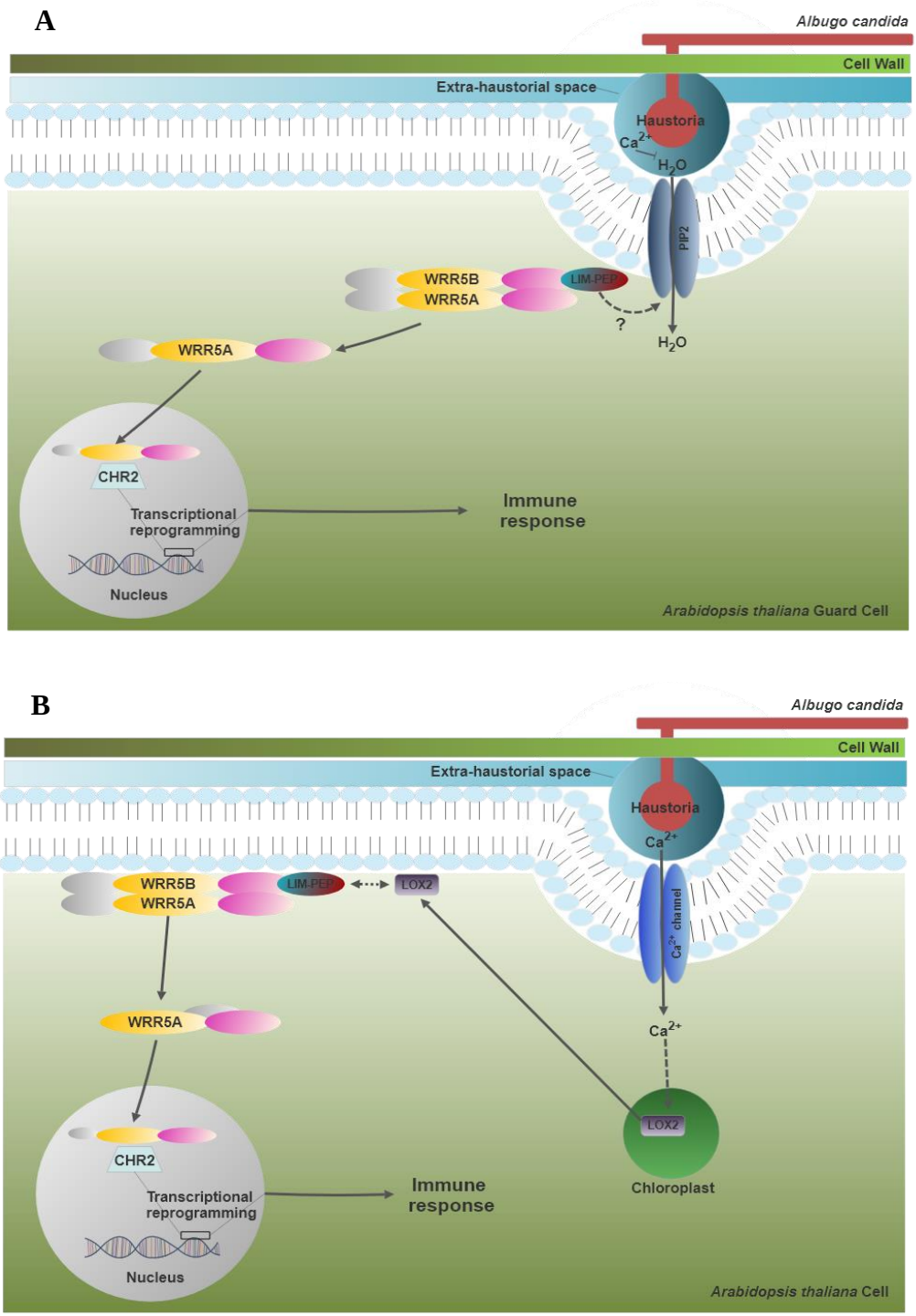


Figure 4.10: PIP2 and LOX2 models of hypothesised *WRR5*A and *WRR5B* immune signalling pathways

Models of the potential WRR5A and WRR5B mechanisms employed to sense an infection by *Albugo candida* isolate AcEM2 in *Arabidopsis thaliana* ecotype Col-0. A - Model of the potential WRR5A and WRR5B mechanism employed to sense an infection by *A. candida* isolate AcEM2 in *A. thaliana* ecotype Col-0 showing interaction with aquaporin PIP2. B – Model of the potentially model of WRR5 mediated immunity, whereby the WRR5 complex interacts with Lipozygenase LOX2 to sense the presence of *A. candida* isolate ACEM2.

Chapter 5

Insights into *WRR7* mediated immunity

Introduction

CNLs are present in both monocots and dicots and are the predominant class of NLRs triggering ETI responses in monocots (Wróblewski et al., 2018). CNLs can be subdivided into distinct classes based on motifs and domains that are present in their CC domain, these include the conserved EDVID motif, the RPW8 domain as well as the traditional coiled coil domain. These three groups are widely distributed throughout plant lineages and in some families the CC domain has evolved family specific sub-classes such as the I2-like and SD-CC classes found in the Solanaceae (Bentham, A.R. et al., 2018). Intriguingly, the Brassicaceae family contains one of the highest TNL:CNL ratios, and CNLs seemingly play a reduced role in this family (Borrelli et al., 2018; Van de Weyer et al., 2019). However, CNLs still make up ~30% of the *A. thaliana* pan-NLRome and therefore play a significant role in the defence of *A. thaliana* plants (Van de Weyer et al., 2019). To date, very little has been identified about the common signalling mechanisms employed by any of the CNL families (Bentham, A.R. et al., 2018). The signalling components that have been identified such as NDR1 are limited to a small amount of CNL systems, but these do not seem to have broader involvement in CNL signalling (Day, Dahlbeck and Staskawicz, 2006a; Knepper, Savory and Day, 2011). A lot of our current understanding of the plant immune system comes from studying the model plant *A. thaliana*, however even in this model very little known about CNL signalling.

WRR7 is an atypical resistance gene from the CC_R family of resistance genes, it consists of an N-terminal RPW8 domain, a partial NB-ARC domain (with no functional P-loop), an integrated LIM-Peptidase domain at its C-terminus but lacks a leucine rich repeat (LRR) domain. *WRR7* confers resistance in *A. thaliana* ecotype Col-0 against *A. candida* race 4 isolate AcEM2 and to our knowledge, does not require a partner NLR unlike in the *WRR5A* and *WRR5B* system discussed in chapter 4. Therefore, studying the mechanistic function of *WRR7* could shed some light on the signalling and regulation of CNL genes in *A. thaliana*.

In this chapter, we performed a forward genetic screen using the *A. thaliana Col-eds1.2* line to identify functionally important sites in *WRR7* and to discover any unknown genes involved in the regulation of *WRR7* or the downstream defence mechanism that *WRR7* employs to stimulate immunity. We identified a mutation in the first α-helix of the *WRR7* RPW8 domain which abolished the immune function of *WRR7*. In addition, we identified several proteins, some of which have not previously been implicated in plant immune signalling. These include Calmodulin binding transcriptional activator (CAMTA2), Chromatin remodelling complex 4 (CHR4), Mitogen-activated protein kinase (MAP3Kδ4) and MOS4-associated complex 7 (MAC7).

Due to the severity of the defence response triggered by NLRs they are tightly controlled by transcriptional regulators to avoid any non-specific gene activation. Therefore, any signals upregulating NLR genes have to be highly specific and the signals that elicit activation of NLRs following pathogen invasion are still being determined. However, there are several key signalling molecules that are altered following pathogen invasion which could trigger expressional upregulation of NLRs such as *WRR7*. One of the most important signalling molecules in plants is Ca^{2+}, our forward genetic screen identified a calcium responsive transcription factor CAMTA2 as a component of WRR7 mediated immunity. Alteration in Ca^{2+} concentrations in different cellular locations trigger both biotic and abiotic stress responses (Kudla et al., 2018). Plants mobilise Ca^{2+} from cellular compartments that act as Ca^{2+} sources during stress, such as the cell wall, vacuole and endoplasmic reticulum to rapidly alter Ca^{2+} concentrations in other cellular compartments such as the cytosol and nucleus (Dodd, Kudla and Sanders, 2010). The range of responses caused by changes in cellular Ca^{2+} means that there are multiple Ca^{2+} signalling stimuli that cause different responses based on the direction, amplitude and spatiotemporal nature of the change in calcium concentration. Ca^{2+} signatures are then decoded by proteins that can perceive changes in Ca^{2+} homeostasis by the binding of Ca^{2+} to calcium binding domains such as EF hands (Dodd, Kudla and Sanders, 2010). The decoding proteins fall into four distinct classes: Calmodulins (CaM), Calmodulin like proteins (CaM-like), Calcium dependent protein kinases (CPKs) and Calcineurin b-like proteins (Ranty, Aldon and Galaud, 2006; Kudla, Batistič and Hashimoto, 2010; Poovaiah et al., 2013; Yuan et al., 2017; La Verde, Dominici and Astegno,

2018; Yip Delormel and Boudsocq, 2019). Biotic interactions between plants and microorganisms are already known to be highly dependent on calcium signalling for example the formation of mutualistic symbioses with arbuscular mycorrhizal fungi are reliant on the advent of nuclear Ca^{2+} spiking events that allow plants to distinguish between pathogenic and mutualistic interactors (MacLean, Bravo and Harrison, 2017). Not only has calcium signalling been shown to play a role in the establishment of mutualisms but Ca^{2+} homeostasis has been shown to be perturbed during pathogenic plant-microbe interactions. During MTI, recognition of microbe associated molecular patterns (MAMPs) at the cell surface are translated into defence responses by an influx of calcium across the plasma membrane into the cytosol, this signal is then decoded by calcium binding proteins which activate downstream signalling (Yuan et al., 2017; Hander et al., 2019). Calcium signalling has not only been linked to MTI in immune signalling but calcium responsive downstream regulators such as the Calmodulin binding transcription activators (CAMTAs) have been identified to play a role in regulating ETI responses through the repression of key TNL downstream signalling gene EDS1 as well as genes that induce salicylic acid responses (Du, Liqun et al., 2009; Kim et al., 2017). One TN protein, TN2, has already been shown to specifically interact with and require CPK5 to cause ETI responses when its guardee EXO70B1 is disrupted by powdery mildew (*Golovinomyces cichoracearum*) effectors in *A. thaliana* (Zhao, T. et al., 2015; Liu et al., 2017). It is therefore highly plausible that a Ca^{2+} signalling event is responsible for the upregulation of *WRR7* following *A. candida* infection.

The expression of NLR genes is highly regulated due to their ability to stimulate autoimmune responses if activated incorrectly (Li, X. et al., 2001). Therefore, understanding the regulation of NLR expression is important in understanding how these genes operate. To understand the regulation of NLR genes, forward genetic screens have been utilised to identify genes able to supress autoimmune phenotypes in lines with amplified NLR activity. For example, the autoimmune mutant of Suppressor of *npr1-1* (*snc1*) constitutively activates defence responses and has been the basis of several screens (Li, X. et al., 2001; Monaghan et al., 2010; Xu et al., 2012b). These screens identified mutations in genes that are modifiers of *snc1* (*mos)* that rescue the autoimmune phenotype by regulating the transcript levels of *snc1* e.g. *mos4* and *mos12* are implicated in the regulation of the correct splicing of the *SNC1*

and *RPP4* transcripts through the MOS4 associated complex (MAC), *MOS9* regulates the methylation state of *SNC1* and *MOS3*, *MOS6* and *MOS7* are all involved in the nucleo-cytoplasmic trafficking of mRNA transcripts (Monaghan et al., 2010; Xu et al., 2012b; Xia, S. et al., 2013). The MAC complex is a conserved complex in eukaryotes that regulates transcript splicing and its homologs in mammals (the Nineteen complex) and yeast (Prp19 complex) have been extensively studied (Jia et al., 2017). Our screen revealed that MAC7, a MOS4 associated protein is important in the regulation of *WRR7* transcripts following AcEM2 infection.

In addition to splicing regulation by MAC7, we also identified a chromatin remodelling protein (CHR4) as being involved in the regulation of *WRR7* transcription. DNA is packaged into higher order structures in the nucleus, this involves the wrapping of DNA around histone proteins into nucleosomes that are separated by linker strands of DNA. Changes induced in the chromatin structure lead to the regulation of large suits of genes, for example Polycomb group proteins trimethylate lysine (K) residues in histones leading to the repression of particular target genes and are key regulators of developmental programmes in eukaryotic organisms (Hennig and Derkacheva, 2009). The post-translational modifications imposed by epigenetic regulators are highly specific, for example Polycomb repressive complex 2 (PRC2) acts specifically on lysine 27 of histone 3 (H3K27) to cause repression of target genes (Hennig and Derkacheva, 2009). As well as polycomb group proteins, chromatin remodelling complexes are involved in the epigenetic regulation of genes by repositioning nucleosomes or post-translationally modifying histones to facilitate the binding or removal of transcription factors (Aalfs and Kingston, 2000; Ha, 2013). The chromatin remodelling factors fall into four discrete subfamilies: Chromodomain-helicase-DNA (CHD), switch/sucrose-non-fermenting (SWI/SNF), imitation switch (ISWI) and the inositol recruiting 80 (INO80) families (Tyagi et al., 2016). CHR4 is part of the CHD family of chromatin remodellers. The CHD family is further divided into 3 sub-families based on the presence or absence of a DNA binding domain (present in subfamily I), a plant homeo domain (PHD) which is present in subfamily II and Brahma and kismet (BRM) and SANT domains that are present in subfamily III (Marfella and Imbalzano, 2007; Gentry and Hennig, 2014). CHD chromatin remodelling factors

have already been implicated in regulating immune responses in plants, including regulation of *SNC1* (Zou et al., 2017; Ramirez-Prado et al., 2018). In *A. thaliana*, the CHD protein subfamily consists of CHR4, CHR5, CHR7 and PICKLE (Hu, Lai and Zhu, 2014). CHR4, PICKLE and CHR7 form part of the CHD subfamily II and are most closely related to mammalian Chd3 whereas CHR5 is the sole member of the Chd1 subfamily in *A. thaliana* (Hu, Lai and Zhu, 2014). Members of this subfamily have been associated with the specific trimethylation of histone H3 lysine 27 (H3K27me3) and act in an antagonistic manner to PRC2 Polycomb group proteins (Zhang, H. et al., 2012; Carter et al., 2018; Jing, Lin and Guo, 2019). Furthermore, it was recently identified that CHR4 regulates both H3K27me3 and H3K4me3 marks and that the *WRR7* loci was hypermethylated for both marks in the *chr4-2* mutant (Sang et al., 2020).

Our forward genetic screen also identified a Map kinase kinase kinase δ4 (MAP3Kδ4) that is involved in the downstream WRR7 defence response. MAP3Kδ4 has been shown to interact with proteins that regulate Abscicic acid (ABA) signalling including protein phosphatase 2Cs (PP2Cs) and ABA-responsive element binding transcription factors (ABFs) and is thought to be a negative regulator of ABA responses as well as being implicated in osmotic stress responses (Lumba et al., 2014).

As well as identifying novel proteins involved in the WRR7 immune pathway, we also identified a structural point mutation in the RPW8 domain of WRR7 that was able to abolish WRR7 mediated resistance to *A. candida*. Recent structural studies on the CNL HOPZ-activated disease resistance 1 (ZAR1), revealed that proceeding activation by the interchange of ADP for ATP from the P-loop of the NB-ARC domain following infection by *Xanthomonas campestris pv. campestris* the N-terminal domains of ZAR1 proteins homodimerized, forming a pentameric complex with a protruding ‘funnel’ of α-helices thought to form a pore capable of perforating the plasma membrane (Wang, J., Wang, et al., 2019). This structural insight provides a compelling mechanism by which resistance genes such as WRR7 may be able to signal defence responses by creating an ion channel across a membrane following recognition of a pathogen leading to downstream defence activation.

To date, little is known about how CC_R or CN_R proteins signal following pathogen infection. However, increasing interest is being paid to these families of resistance proteins due to the recently identified role of two CC_R families: Activated disease resistance 1 (ADR1s) and N requirement gene (NRGs) in performing downstream signalling functions for other resistance proteins such as Roq1, RPP1 and WRR4 (Bonardi et al., 2011; Dong, O.X. et al., 2016; Castel et al., 2018; Qi et al., 2018; Lapin et al., 2019). ADR1s and NRGs are being shown to act downstream of the EDS1-SAG101 or EDS1-PAD4 complexes in TNL mediated immunity (Gantner et al., 2019; Lapin et al., 2019). Sequence analysis of CNLs and CN proteins encoded by the *A. thaliana* Col-0 genome revealed that WRR7 forms a clade with the NRG family proteins NRG1.1 and NRG1.2 (Wróblewski et al., 2018). Interestingly, we found that WRR7 enacts defence against *A. candida* independently of other NLRs and to date doesn't have any known role as a helper NLR. Therefore, WRR7 may have originally performed a similar role to the NRG proteins but has since neofunctionalised. This raises the interesting question of how this CC_R protein enacts defence and whether, by studying this protein, we could glean any functional information on how other CC_Rs upregulate downstream defences.

Results

WRR7 is upregulated in *Arabidopsis thaliana* ecotypes following infection by *Albugo candida* isolates.

The *A. thaliana* ecotype Ws-2 is susceptible to *A. candida* isolate AcEM2 (Fig 3.3). However, a copy of *WRR7* is present in the Ws-2 genome that contains two SNPs in exons 2 and 7 as well as one SNP in the promoter region compared to the Col-0 *WRR7* allele. Only 1 of these SNPs causes a non-synonymous M528I amino acid substitution (Fig 5.1), both Methionine and Isoleucine share similar amino acid properties but cause small changes in hydrophobicity (Ohmura et al., 2001). Therefore, we predicted that these changes alone were unlikely to be the reason why *WRR7* is non-functional in the Ws-2 ecotype and that it was likely that the misexpression of the gene was the cause of AcEM2 susceptibility in Ws-2. We also wanted to test the specificity of the *WRR7* response to different *A. candida* races to determine whether *WRR7* is a specific response to AcEM2 infection or whether it can respond to a broader range of *A. candida* races.

Therefore, we performed a RT-qPCR analysis of the susceptible *A. thaliana* line Ws-*eds1* following inoculation with *A. candida* race 4 isolate AcEM2 as well as the race two isolate AcBJ12 (Fig 5.1). We also performed RT-qPCR analysis with the susceptible accession Ws-*eds1* and compared the expression of *WRR7* to the resistant Col-*eds1.2* accession following inoculation with AcEM2 only (Fig 5.1). This allowed us to determine whether *WRR7* is upregulated following AcEM2 infection in both susceptible and resistant interactions. Surprisingly, we found that *WRR7* is significantly up-regulated (log_2 Fold change >1) in susceptible line Ws-*eds1* following AcEM2 infection (Fig 5.1). We also found that *WRR7* expression was also significantly up-regulated in Ws-*eds1* plants challenged with *A. candida* isolate AcBJ12 (Fig 5.1). This showed that *WRR7* is upregulated in susceptible interactions and therefore either the M528I point mutation can inhibit immune activation by *WRR7* or a downstream signaling component of *WRR7* mediated immunity is non-functional in Ws-2 plants. *WRR7* expression was also significantly upregulated after 3 and 5 dpi with AcEM2 in Ws-*eds1* and Col-*eds1.2* plants and showed greater expression in Ws-*eds1* compared to Col-*eds1.2* (Fig 5.1). This increased expression in Ws-*eds1* plants was likely due to a greater number of cells being infected with AcEM2 due to the susceptible interaction.

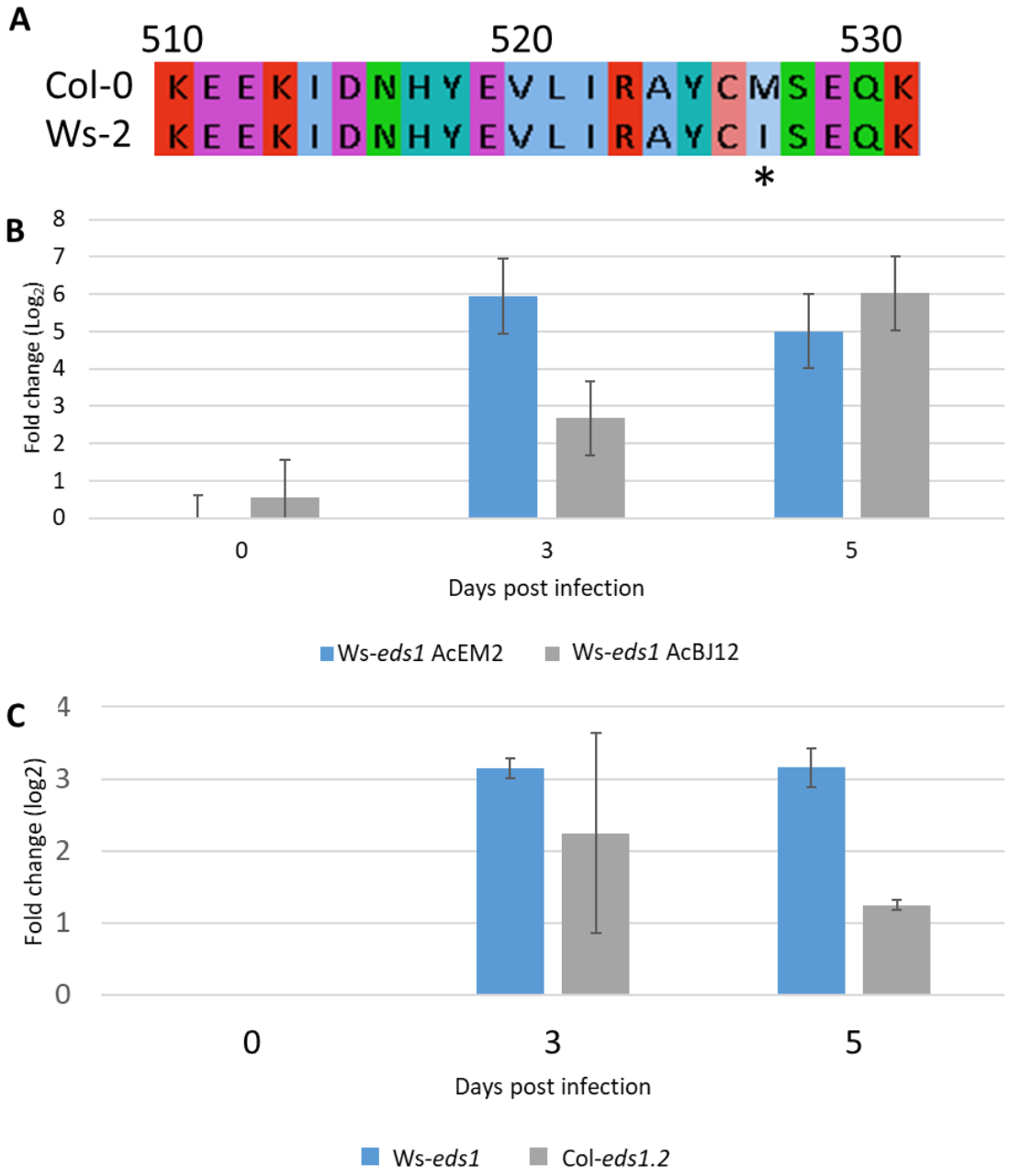


Figure 5.1: *WRR7* expression increases in susceptible and resistant interactions with *Albugo candida* isolates in *Arabidopsis thaliana*
A - Alignment of WRR7 amino acid sequence between *Arabidopsis thaliana* ecotypes Col-0 and Ws-2 using Muscle software (Edgar, 2004). B and C - Bar charts showing the log2 fold change of *WRR7* expression in B) *A. thaliana* line Ws-eds1 0, 3, and 5 dpi with *A. candida* isolates AcEM2 and AcBJ12 and C) *A. thaliana* lines Ws-*eds1* and Col-*eds1.2* 0, 3, and 5 dpi with *A. candida* isolate AcEM2. *WRR7* expression was analysed using the ΔΔCT method, expression was normalised against housekeeping gene *PP2AA3* and expression levels of AcEM2 infected samples were compared against the expression levels of mock inoculated plants of the same age. Error bars represent standard error.

RNA sequence analysis reveals genes regulated by CAMTA transcription factors during colonisation of *A. thaliana* leaves by *A. candida*

Plants respond to changes in their environment by altering the expression of genes involved in responding to external stimuli. Changes to environmental stimuli lead to large transcriptomic alterations that are controlled by particular transcription factors. The expression of *WRR7* increases following infection by *A. candida* (Fig 5.1) and this change in expression is likely due to transcriptomic control of *WRR7* by a particular transcription factor that responds to *A. candida* infection prior to the ETI response. *A. candida* suppresses plant immunity during infection, facilitating colonisation of the host by non-host pathogens (Cooper et al., 2008; Prince et al., 2017) reducing the amount of defence related genes that are upregulated following colonization. Therefore, to determine whether any specific transcription factors control developmental and immune responses in an interaction between *A. thaliana* and *A. candida* prior to an ETI response we performed RNA sequence (RNA seq) analysis on the immunocompromised Col-*eds1.2-wrr7* line.

To detect any transcriptomic changes that occurred during infection we performed a time course experiment whereby three biological replicates were collected at 2, 4 and 6 dpi for pathogen treated and mock (water) treated samples. To test whether the biological samples produced similar results we performed a principle component and hierarchical cluster analysis for each replicate (Fig 5.2). Our data suggest that all three biological replicates for each time point and treatment were highly similar to one another and distinctly different from the other time points and treatments (Fig 5.2). Although, the 2-dpi pathogen treated samples clustered more closely with water treated samples than 4 dpi and 6 dpi samples (Fig 5.2). Therefore, this suggests that there is a delayed response of at least 2 days to infection by *A. candida*. After confirming that our RNA seq sample replicates were similar (Fig 5.2). We determined how many differentially expressed genes (DEGs) were present at each time point by comparing the AcEM2 treated samples to water treated samples (Fig 5.3). We found that there were only 422 DEGs 2 dpi with AcEM2 treated samples compared to 3532 DEGs at 4 and 3655 DEGs at 6 dpi with AcEM2.

To find any genes that could be controlled by specific transcription factors we performed a K-means cluster analysis to group genes with similar expression profiles. To do this, genes that showed differential expression at at least one time

point were selected and the biological replicates for each gene were summarised by generating Z-score values, the Z-scores were then used to perform a K-means cluster analysis based on differentially expressed genes in iDEP 9.0 (Ge, Son and Yao, 2018). We identified 10 clusters of genes showing similar expression profiles after infection by AcEM2 (Fig 5.4). The promoter regions (500 bp upstream of the start codon) of the genes identified in each cluster were then analysed for the presence of any common transcription factor binding motifs using Pscan (Zambelli, Pesole and Pavesi, 2009) to determine the most likely transcription factors regulating each cluster after infection with AcEM2 (Table 5.1).

We found that two clusters of genes, cluster H and J, were significantly enriched for CAMTA transcription factor binding sites in their promoter regions (Table 5.1). CAMTA transcription factors are associated with a conserved regulatory promoter CG(C/T)GT motif that is found in the promoter sequence of genes that they regulate (Fig 5.5). The genes in clusters H and J showed decreased transcript levels at 2 dpi followed by increased transcript levels after 4 dpi and the transcript levels in genes in both clusters decreased again at 6 dpi. Interestingly, *WRR7* contains a T-DNA insert (derived from the SALK_068218 line) in the Col-*eds1.2-wrr7* line that is located in the intron between exons four and five. Therefore, in the RNA-Seq analysis we still obtained reads from *WRR7* corresponding to the first four exons. The reads for this segment of *WRR7* showed differential expression similar to the expression profile observed during a resistant reaction (Fig 5.1). The expression profile generated from the *WRR7* reads was located in cluster J, a cluster that shows significant enrichment for CAMTA transcription factors and the *WRR7* promoter region contains a CGCGT CAMTA binding motif 93 bp upstream of the start codon (Fig 5.5). We also found that *DAR7* (see chapter 6) was also part of cluster J and contains a CAMTA binding motif in its promoter region (Fig 5.4). Interestingly, other defence related genes that are involved in SA biosynthesis showed contrasting expression, for example *EDS5* was significantly downregulated and was part of Cluster D, whereas *CBP60g* was upregulated and clustered into cluster J that is enriched for regulation by CAMTA transcription factors (Fig 5.4).

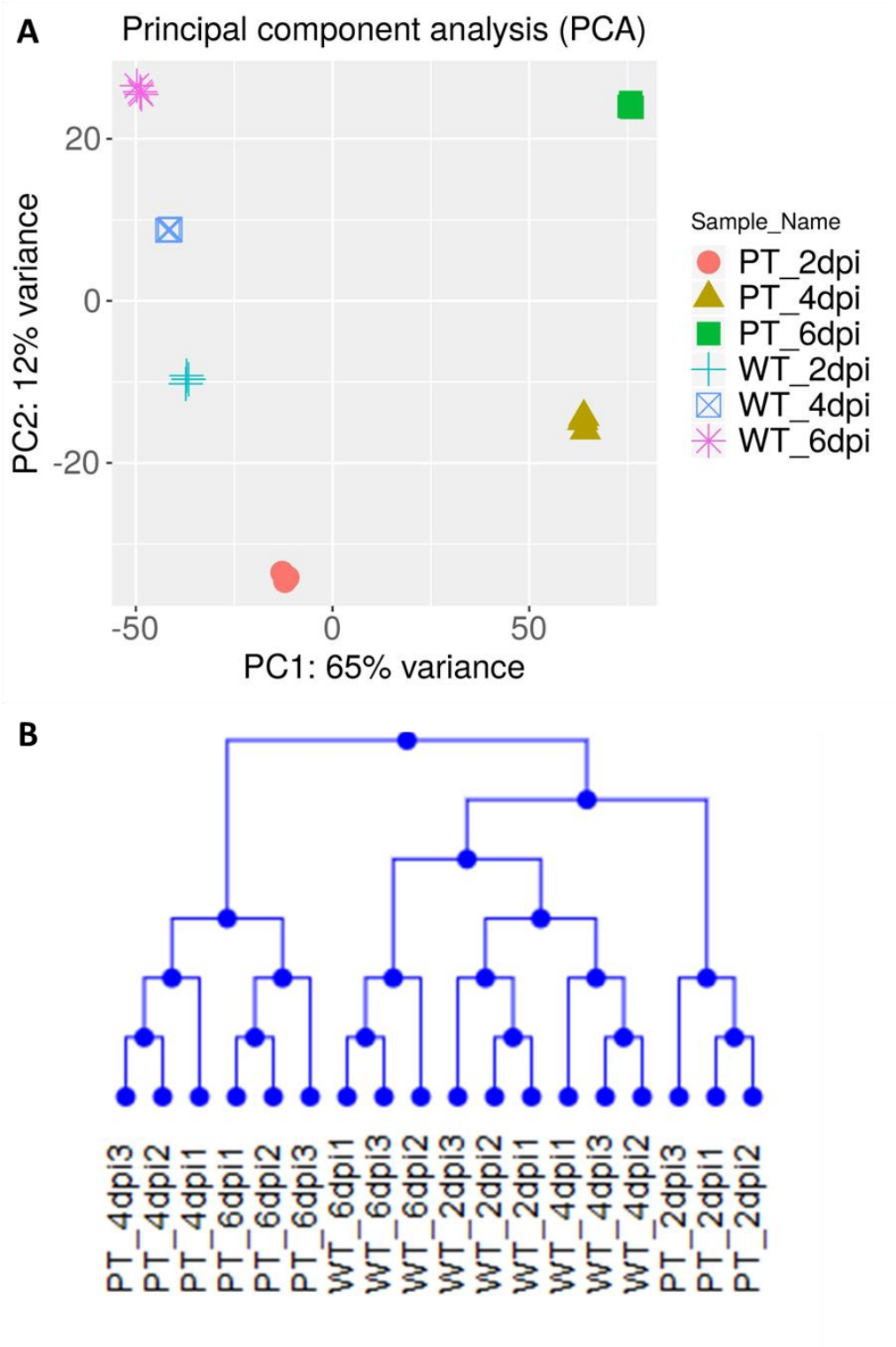


Figure 5.2:Principle component and Hierarchical cluster analysis show that RNA seq replicates at each treatment and timepoint are highly similar
Analysis of RNA sequencing sample replicates of Col-*eds1.2-wrr7* plants infected with *Albugo candida* isolate AcEM2 2, 4- and 6-days post infection (dpi) in pathogen treated (PT) and water treated (WT) samples. A- Scatter plot of principle component analysis. B- Dendrogram of hierarchical cluster analysis results.

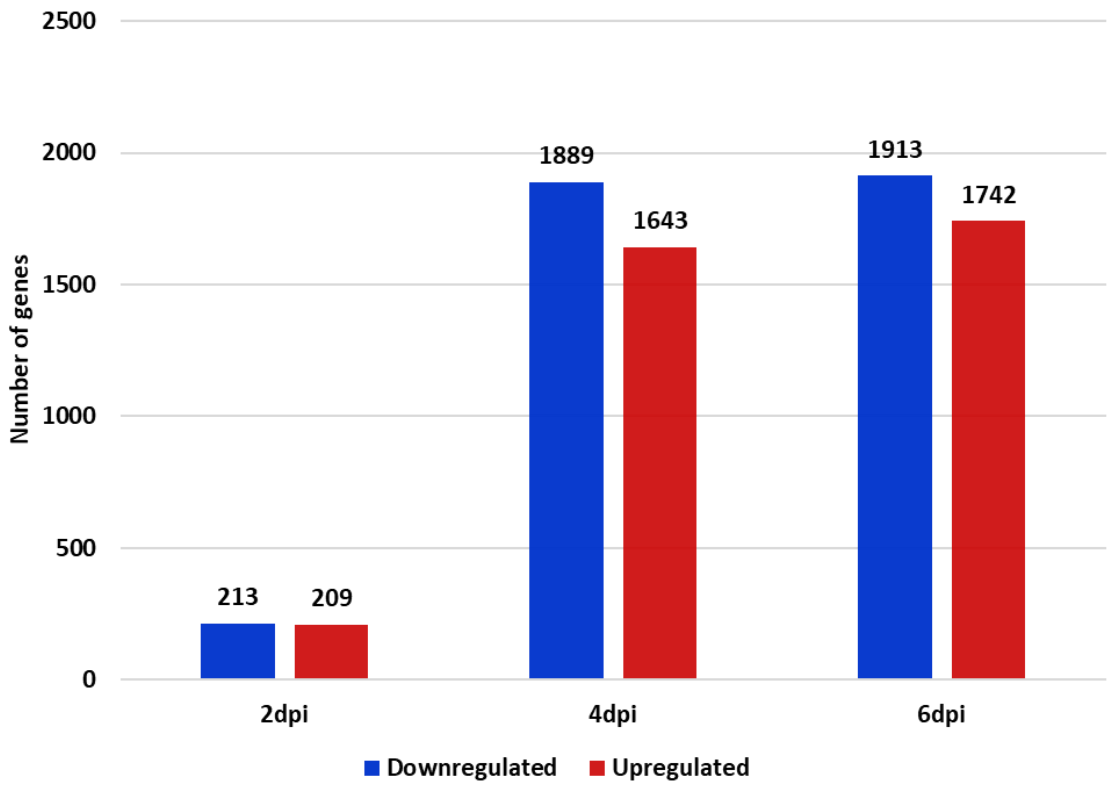


Figure 5.3: Number of differentially expressed genes after infection by *Albugo candida*
Bar chart showing the number of significantly (Log_2 fold change >1) upregulated or downregulated genes in Col-*eds1.2-wrr7* plants infected with *Albugo candida* isolate AcEM2 2, 4- and 6-days post infection (dpi) compared to Col-*eds1.2-wrr7* plants mock inoculated with water at the same time point..

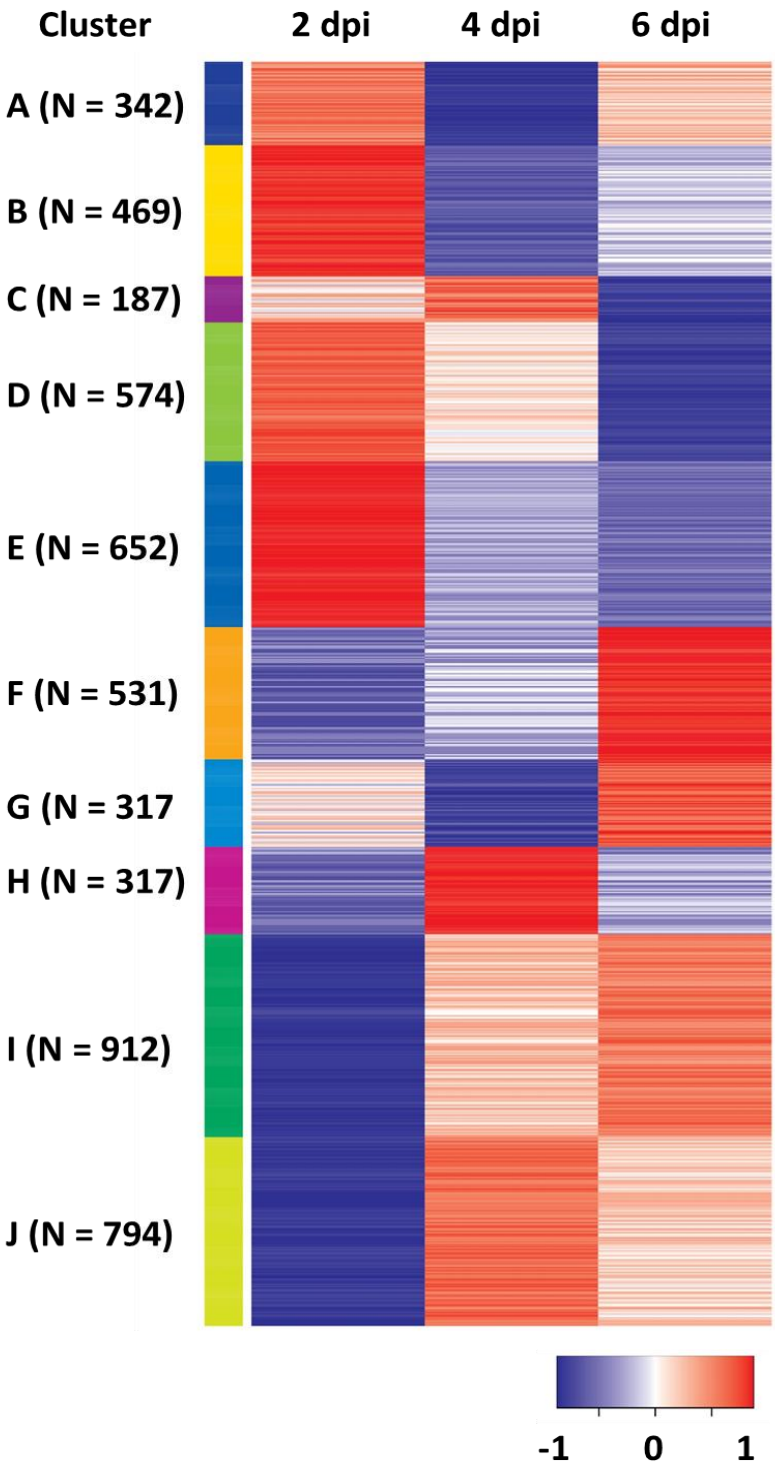


Figure 5.4: K-means cluster analysis of RNA sequence data from Col-*eds1.2-wrr7* plants identifies 10 clusters of similarly expressed genes following infection with *Albugo candida* isolate AcEM2

Heat map of differentially expressed genes (DEGs) in *Arabidopsis thaliana* Col-*eds1.2-wrr7* plants after infection by *Albugo candida* isolate AcEM2 2, 4- and 6-days post infection (dpi) based on the Z-scores of the three combined replicates for each time point. DEGs were separated into 10 clusters based on their expression profiles by K-means cluster analyses using iDEP software (Ge, Son and Yao, 2018).

Table 5.1: Transcription factor enrichment analysis reveals that CAMTA transcription factors control the upregulation of genes following *Albugo candida* infection

Top 5 candidate transcription factors predicted to regulate genes in each cluster identified by K-means cluster analysis. Transcription factor binding analysis performed using Pscan software (Zambelli, Pesole and Pavesi, 2009).

Cluster	Transcription factor	P-value
A	AREB3	4.9947E-06
A	At5g08520	2.0008E-05
A	bZIP16	2.289E-05
A	At1g19000	2.6071E-05
A	TCP23	3.2363E-05
B	MYB3R4	2.8684E-17
B	MYB3R1	3.9906E-16
B	MYB3R5	5.5775E-16
B	bZIP68	1.1623E-09
B	bZIP16	2.3931E-09
C	BZR2	1.2946E-10
C	PIF7	1.8745E-10
C	UNE10	2.0458E-10
C	AT4G18890	2.7532E-10
C	PIF1	4.6893E-10
D	ATHB-6	2.4379E-11
D	ATHB-51	5.4746E-11
D	GT3a	1.6521E-10
D	HAT5	9.5809E-10
D	ATHB13	1.3566E-09
E	ABF2	6.7009E-15
E	BEE2	3.3531E-14
E	BIM3	4.217E-14
E	BIM2	5.7099E-14
E	bHLH31	5.1247E-13
F	At1g19000	1.4458E-08
F	At5g08520	6.7832E-08
F	At5g08330	8.1192E-08
F	At1g72010	1.5009E-07
F	At5g58900	2.2122E-07
G	AT5G56840	4.9957E-14
G	At1g19000	5.752E-14
G	At5g58900	6.5999E-14
G	At5g47390	2.9498E-13
G	AT3G10580	9.3584E-13
H	CAMTA1	8.9952E-11
H	CAMTA2	1.5165E-10
H	CAMTA3	1.2144E-08
H	bZIP68	8.6866E-07

H	bZIP16	1.4423E-06
I	AT2G28810	4.0707E-18
I	AT5G66940	1.2041E-17
I	OBP1	2.5694E-17
I	AT5G02460	8.9585E-17
I	OBP3	1.0683E-15
J	CAMTA1	5.3661E-19
J	CAMTA2	8.2777E-19
J	CAMTA3	6.4362E-16
J	ABI5	1.6437E-12
J	HY5	1.2465E-11

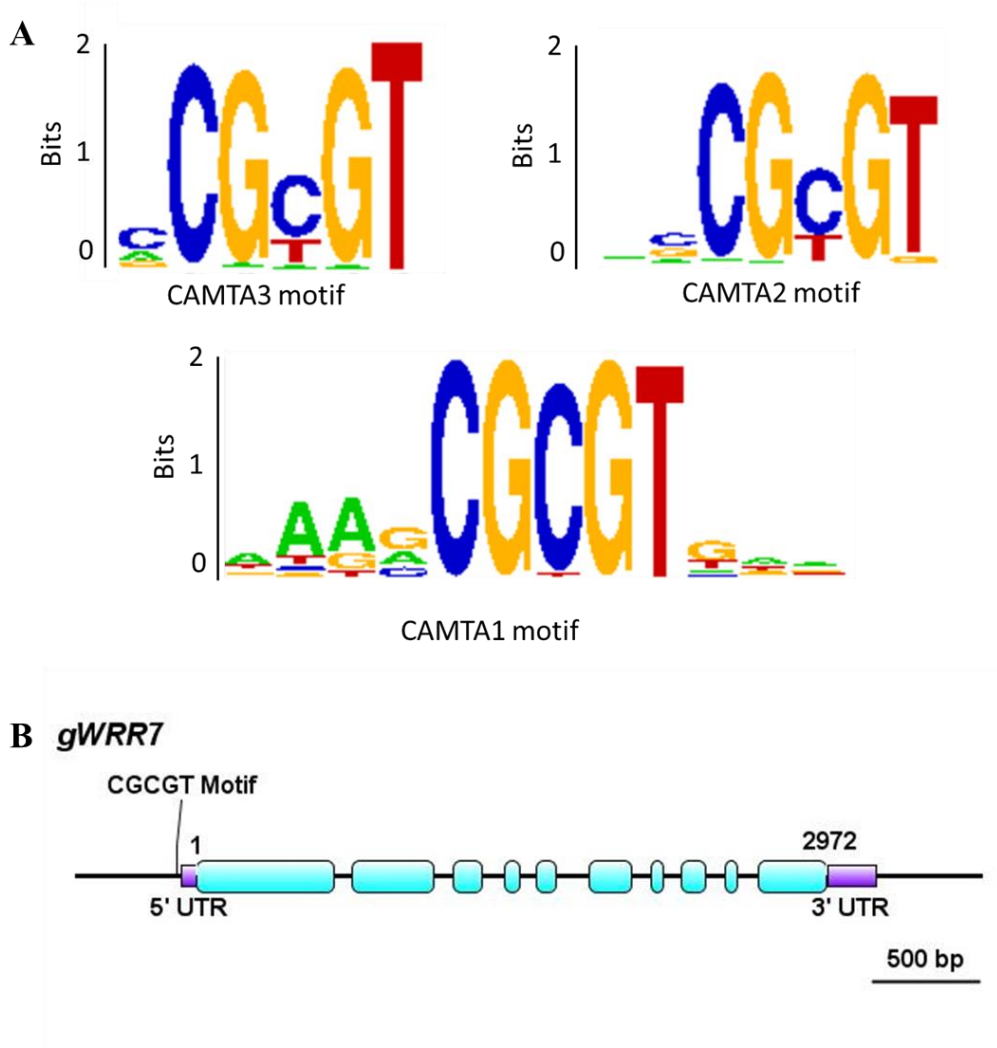


Figure 5.5: CAMTA transcription factor binding motifs
A- Logos of transcription factor binding motifs associated with CAMTA1, CAMTA2 and CAMTA3 as predicted by Jaspar 2018 (Khan et al., 2018). B- *WRR7* nucleotide structure showing the location of the CAMTA binding motif. Gene is numbered from the start codon to the stop codon, purple boxes represent untranslated regions (UTR), blue boxes represent exons and the scale bar represents 500 base pairs).

Identification of genes involved in *WRR7* immune signalling via EMS mutagenesis

In chapter 3 figures 3.8 and 3.10, I showed that *WRR7* confers resistance to *A. candida* in *A. thaliana* ecotype Col-0. *WRR7* is a CC_R type resistance gene. We currently have a very poor understanding of how CC_Rs containing RPW8 domains at their N-termini cause resistance. The Col-0 genome contains not only CN_R WRR7 but also TNLs WRR4A, WRR5A and WRR5B that can cause resistance to *A. candida* race 4 isolate AcEM2. TNLs require lipase-like protein EDS1 for their immune function. Therefore, we performed an EMS screen on *Col-eds1.2* to identify any proteins that are involved in *WRR7* mediated immunity that may also be involved in immunity caused by other CC_Rs or be involved in the downstream signalling of helper CC_Rs such as the ADR1 and NRG protein families (Fig 5.6). After mutagenesis, mutated (M_1) seedlings were grown and M_2 seeds were harvested from each individual M_1 line. M_2 seed pools were sown and 2-week-old seedlings were screened with *A. candida* isolate AcEM2, seedlings displaying different levels of susceptibility were selected and surviving plants were then grown and re-tested with AcEM2 after 5 weeks of growth to confirm the susceptible phenotype. Plants maintaining a susceptible phenotype after the second screen were then selfed, generating M_3 populations from each AcEM2 susceptible progenitor line. Seeds from multiple M_3, populations derived from each susceptible M_2 mutant, were then sown and challenged with the pathogen to identify a homozygous, susceptible mutant line from each susceptible M_2 mutant. Homozygous susceptible M_3 plants were then backcrossed (BC) to *Col-eds1.2*, grown, and selfed to generate BC F_2 populations that were segregating for AcEM2 susceptibility. Susceptible seedlings (~100) from each BC F_2 population were selected from these populations and combined, DNA was then extracted from this bulked population and sent for whole genome re-sequencing.

From the 191 M_2 seed pools (corresponding to ~200,000 seedlings) we identified 99 susceptible seedlings from 41 of the M_2 seed pools, only 73 of these seedlings were recovered as some of the seedlings were overcome with the pathogen and four lines were male sterile. M_3 seeds from these populations were sown and re-tested with AcEM2, only seven of these lines (from now on referred to as 'EMS' lines) showed a consistent susceptible phenotype and were taken forward for further analysis (Fig

5.6). The susceptibility of these lines varied from EMS 138 which showed green susceptibility i.e. pustule development with no observable cell death to EMS 1 which showed minor pustule development, often only one or two per leaf with extensive cell death (Fig 5.6). Interestingly, we observed pustule development alongside cell death responses in 6/7 of the identified mutants suggesting that these mutants could still activate a partial defence response, but this response was not strong enough to inhibit pathogen colonisation (Fig 5.6).

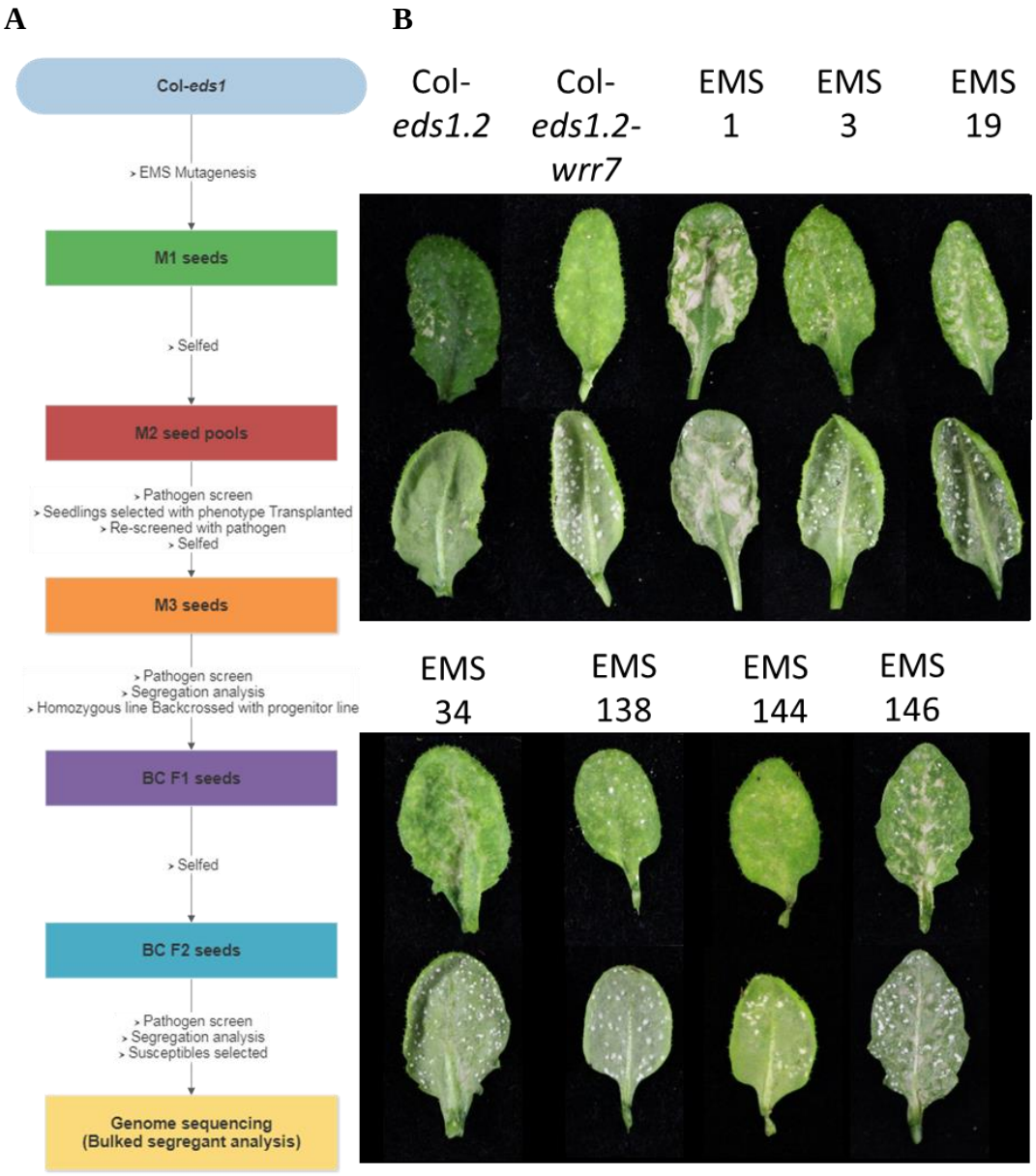


Figure 5.6: Identification of *Arabidopsis thaliana* Col-*eds1.2* plants susceptible to *Albugo candida* isolate AcEM2 after EMS mutagenesis

A- Flow diagram representing the crossing stages performed to identify mutants. B – leaf images of *Arabidopsis thaliana* Col-*eds1.2* M_3 EMS mutant phenotypes 10 days post infection with *Albugo candida* isolate AcEM2

Bulk segregant analysis of susceptible BC F_2 plants reveals novel players in *WRR7* mediated immunity

EMS mutagenesis typically induces single nucleotide polymorphisms (SNPs), resulting in base pair changes from G/C to A/T, due to the alkylation of guanine residues (Sikora et al., 2012). The induced SNPs occur at ~150-300 kb intervals in the genome on average and about 12% of these result in the loss of function of a gene (Bökel, 2008). In the *A. thaliana* genome, we would expect EMS mutagenesis to result in ~1000 such induced mutations per genome at this rate. Identifying which of these SNPs is the causal mutation for any identified phenotype requires a whittling down process that involves several rounds of selfing followed by backcrossing and further rounds of selfing before the candidate causal SNPs can be identified by direct sequencing (Sikora et al., 2012). SNPs most likely to cause the susceptible phenotypes of our EMS mutants were expected to be G/C to A/T mutations causing non-synonymous mutations or truncations of their translated protein. Therefore, candidate SNPs were only considered if they met these criteria.

We identified five candidate mutant genes from our seven EMS mutant lines using the SIMPLE pipeline (Wachsman et al., 2017). Only one of our mutants (EMS 138) was identified to have a mutation in *WRR7* itself (Table 5.2), this was a C to T nucleotide substitution at nucleotide position 23 causing a non-synonymous S8F mutation in the RPW8 domain of WRR7. Candidate genes were determined by the presence of mutant SNPs only in reads from the susceptible mutant genomes and not in the reference genome and that weren't containing the reference nucleotide in any of the reads in the mutant genome. Three of our mutant lines (EMS 3, 19 and 144) contained SNPs in Ca^{2+} responsive transcription factor *CAMTA2* (*AT5G64220*) encoding gene (Table 5.2). EMS 19 contained a SNP in *CAMTA2* at the intron-exon boundary causing a splice variant mutant, whereas EMS 3 and EMS 144 both contained SNPs causing a non-synonymous R69Q mutation (Fig 5.7) and were later shown to be derived from the same M_2 line (Table 5.2). EMS 19 also contained a non-synonymous SNP in *FAAH* (*AT5G64440*), and EMS 144 had no clear candidate based on the mapping alone (Table 5.2), but due to the presence of the *CAMTA2* mutation in all three of these mutant lines *CAMTA2* was selected as the most likely candidate gene in EMS 3, 19 and 144. Sequencing of susceptible BC F_2 population from EMS 1 and SIMPLE pipeline analysis revealed a single candidate SNP, that

was only present in the mutant genome and not the reference genome, this SNP occurred in *MAP3Kδ4* (*AT4G23050*) causing a nonsynonymous point mutation resulting in a D609N amino acid change in the protein (Table 5.2 and Fig 5.7). EMS 146 contained only one candidate SNP causing a non-synonymous mutation in spliceosome associated gene *MAC7* (*AT2G38770*), causing a glycine to serine change at amino acid position 890 in the protein sequence (Table 5.2 and Fig 5.7). EMS 34 did not contain a clear-cut candidate gene, due to contamination of DNA from resistant seedlings in the bulked DNA. However, we determined the most likely SNP to cause the susceptible phenotype observed in EMS 34 to be the G to A SNP (Table 5.2), that causes a P23S non-synonymous mutation (Fig 5.7) in the protein sequence of CHR4 (AT5G44800). This was determined because it was the SNP containing the highest ratio of the mutant allele to the wildtype allele in the mutant compared to the reference genome that caused a non-synonymous point mutation in the protein sequence (Table 5.2).

Table 5.2: Bulk segregant analysis by direct sequencing reveals candidate mutant genes responsible for the susceptible phenotypes in each of our *Alblugo candida* isolate AcEM2 susceptible *Col-eds1.2* mutant lines

Bulk segregant genome analysis results from Backcross F_2 populations of *Arabidopsis thaliana Col-eds1.2* EMS mutants backcrossed to *Col-eds1.2*. Genomes of seedlings susceptible to *A. candida* isolate AcEM2 F_2 were bulked and reads were processed using the SIMPLE pipeline (Wachsman et al., 2017). All candidate single nucleotide polymorphisms (SNPs) were sorted based on their absence from the reference *Col-eds1.2* genome and the absence of the wildtype SNP from the EMS mutant genome and the top results are shown. The primary candidate genes for each EMS line are marked with * and written in bold.

EMS line	ATG number	Reference genome Nucleotide	Mutant genome nucleotide	Effect of mutation	Protein change	Number of mutant nucleotide in reference genome	Number of mutant nucleotide in mutant genome	Number of wildtype nucleotide in reference genome	Number of wildtype nucleotide in mutant genome
1	**AT4G23050 ***	G	A	missense variant	D609N	0	15	20	0
1	AT1G36210	G	A	downstream gene variant		1	4	16	0
1	AT3G44470	G	A	upstream gene variant		1	3	5	0
1	AT3G59690	G	A	upstream gene variant		1	5	27	0
1	AT1G36210	C	T	downstream gene variant		1	4	16	0
3	**AT5G64220 ***	G	A	missense variant	R69Q	0	21	18	0
3	AT3G42206	G	T	upstream gene variant		1	5	4	0
3	AT5G63750	C	T	upstream gene variant		1	21	20	0
3	AT1G24938	C	T	upstream gene variant		2	2	10	0
3	AT2G12815	C	A	upstream gene variant		2	4	9	0

19	AT5G62760	G	A	upstream gene variant		0	27	22	0
19	AT5G63290	G	A	upstream gene variant		0	41	25	0
19	**AT5G64220***	G	A	splice acceptor variant		0	50	21	0
19	AT5G64440	G	A	missense variant	G199E	0	62	28	0
19	AT3G47280	G	A	downstream gene variant		1	4	5	0
19	AT5G64740	G	A	missense variant	G121E	1	51	25	0
34	AT5G39490	G	A	upstream gene variant		1	17	14	0
34	AT3G33058	G	A	noncoding exon variant		2	2	6	0
34	AT3G33080	G	A	downstream gene variant		2	2	5	0
34	AT3G31970	C	T	downstream gene variant		2	4	7	0
34	AT3G31990	C	T	upstream gene variant		2	5	4	0
34	AT3G33080	C	T	downstream gene variant		2	2	5	0
34	AT3G42256	G	A	upstream gene variant		3	2	8	0
34	AT3G42256	G	A	upstream gene variant		3	2	6	0
34	AT3G44470	G	A	upstream gene variant		3	2	4	0
34	**AT5G44800***	G	A	missense variant	P23S	3	17	20	0
34	AT3G32890	C	T	upstream gene variant		3	8	4	0

144	AT3G43302	G	A	downstream gene variant		0	4	8	0
144	AT3G42490	G	A	upstream gene variant		1	3	4	0
144	AT3G31330	G	A	upstream gene variant		2	3	7	0
144	AT3G33065	C	T	upstream gene variant		2	2	6	0
144	AT1G67870	G	A	synonymous variant	T268T	3	8	18	0
144	AT3G32880	G	A	upstream gene variant		3	4	5	0
144	AT3G42256	G	A	upstream gene variant		3	3	8	0
144	AT3G42256	G	A	upstream gene variant		3	3	6	0
144	AT4G06624	G	A	upstream gene variant		3	3	10	0
144	AT3G44390	C	T	upstream gene variant		3	2	8	0
144	AT2G11502	G	A	upstream gene variant		4	5	12	0
144	AT2G37460	G	A	upstream gene variant		4	3	26	0
144	AT3G09840	G	A	missense variant	A146T	4	8	15	0
144	AT3G32090	G	A	upstream gene variant		4	3	13	0
144	AT5G38360	G	A	upstream gene variant		4	10	14	0
144	**AT5G64220***	G	A	missense variant	R69Q	4	13	18	0
146	AT3G43302	G	A	downstream gene variant		0	4	5	0

146	**AT2G38770***	C	T	missense variant	G890S	0	16	20	0
146	AT2G32730	C	T	missense variant	A797V	1	18	27	0
146	AT3G43304	C	T	upstream gene variant		1	5	11	0
146	AT3G31720	G	A	upstream gene variant		2	5	4	0
146	AT1G35560	G	A	upstream gene variant		3	13	27	0

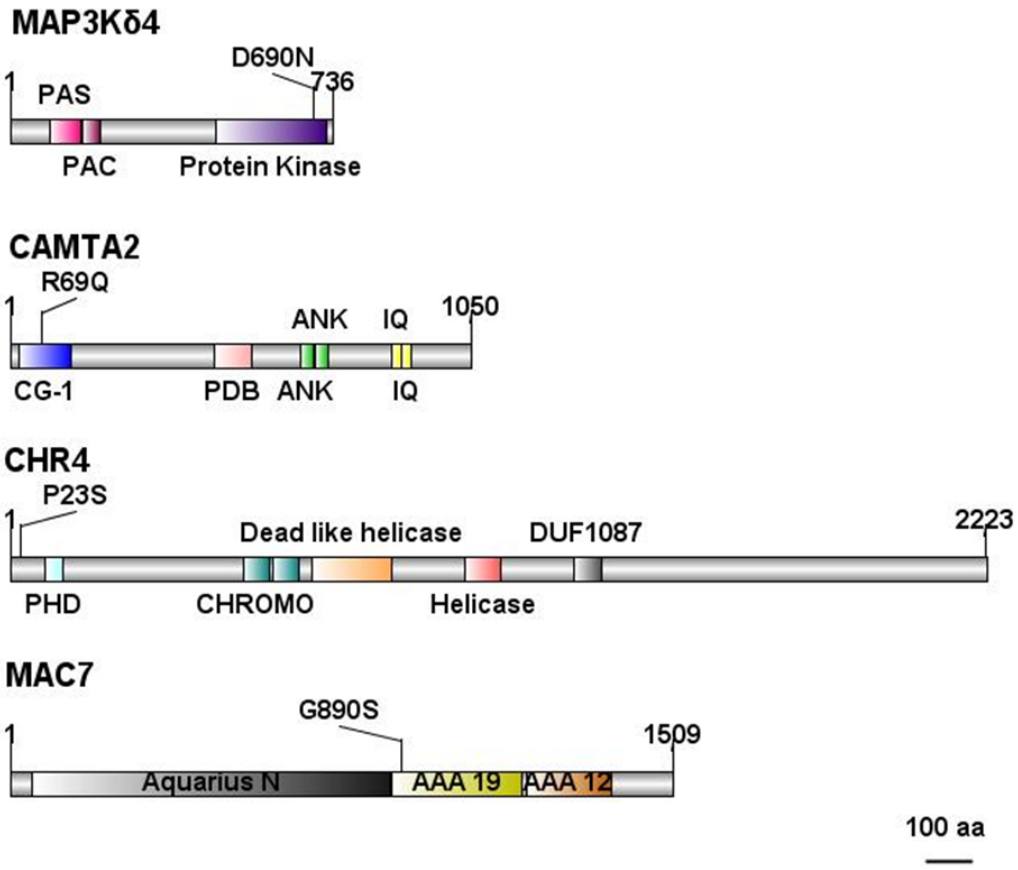


Figure 5.7: Non-synonymous mutation locations in the EMS candidate proteins
Schematic diagram of MAP3Kδ4 (At4G23050), CAMTA2 (At5G64220), CHR4 (At5G44800) and MAC7 (At2G38770) protein domains as predicted by SMART (Letunic, Ivica, Doerks and Bork, 2015). The location of the non-synonymous mutations identified from EMS 1, EMS 3, EMS 34, EMS 144 and EMS 146 are marked in their respective genes. Scale bar represents 100 amino acids.

Cloning and complementation of candidate genes in the EMS lines

The candidate genes identified by the bulk segregant analysis (*CAMTA2*, *CHR4*, *MAP3Kδ4* and *MAC7*) were cloned from Col-0 DNA (*CAMTA2* and *CHR4*) or from JATY clones 52P13 (*MAC7*) or 70J04 (*MAP3Kδ4*) (Fig 5.8). The promoter regions from CAMTA2, CHR4, *MAP3Kδ4* and *MAC7* were cloned 781, 3683, 1496 and 640 bp upstream of the start codons of the gene respectively (encompassing the intergenic region between the gene of interest and the gene directly upstream) and the terminator sequences extended 698, 2862, 1671 and 693 bp beyond the stop codon respectively (Fig 5.8). Once cloned these constructs were transformed into their respective EMS lines (*CAMTA2* into EMS 3, 19 and 144, *CHR4* into EMS 34, *MAP3Kδ4* into EMS 1 and *MAC7* into EMS 146) to determine whether they could recover the susceptibility phenotype imposed on the mutant lines by the mutagenesis.

We found that genomic *CAMTA2* constructs transformed into EMS 19, 3 and 144 were able to recover resistance against *A. candida* isolate AcEM2 in T_1 lines (Fig 5.9). In total, 36 independent T_1 lines of *gCAMTA2* in EMS 19, 15 in EMS 3 and 13 in EMS 144 and of these 25, 13 and 10 showed full resistance respectively (Table 5.3), five non-transgenic seedlings from each of these lines all showed susceptible phenotypes. Therefore, CAMTA2 is a transcription factor that is required for the immune response mediated by *WRR7*. Furthermore, we found that the susceptible phenotype observed in Col-*eds1.2* EMS mutant lines 1, 34 and 146 could be recovered by transformation with *gMAP3Kδ4, CHR4*, and *MAC7* respectively (Fig 5.10 and Table 5.3). We were able to recover 7 independent T_1 lines from EMS 34 and EMS 146 transformed with *gCHR4* and *gMAC7* respectively, with all 7 EMS 146 lines showing resistance to AcEM2 and 5 of the 7 EMS 34 lines compared to 4/4 non-transgenic seedlings that were susceptible to AcEM2 from both of these lines (Table 5.3). We were only able to recover 4 independent EMS 1 T_1 lines that were successfully transformed with *gMAP3Kδ4*, all of which displayed full resistance to AcEM2 compared with 4 non-transgenic lines that were all susceptible (Table 5.3). However, a larger sample size is needed to confirm whether this gene can fully complement the susceptible phenotype of EMS 1.

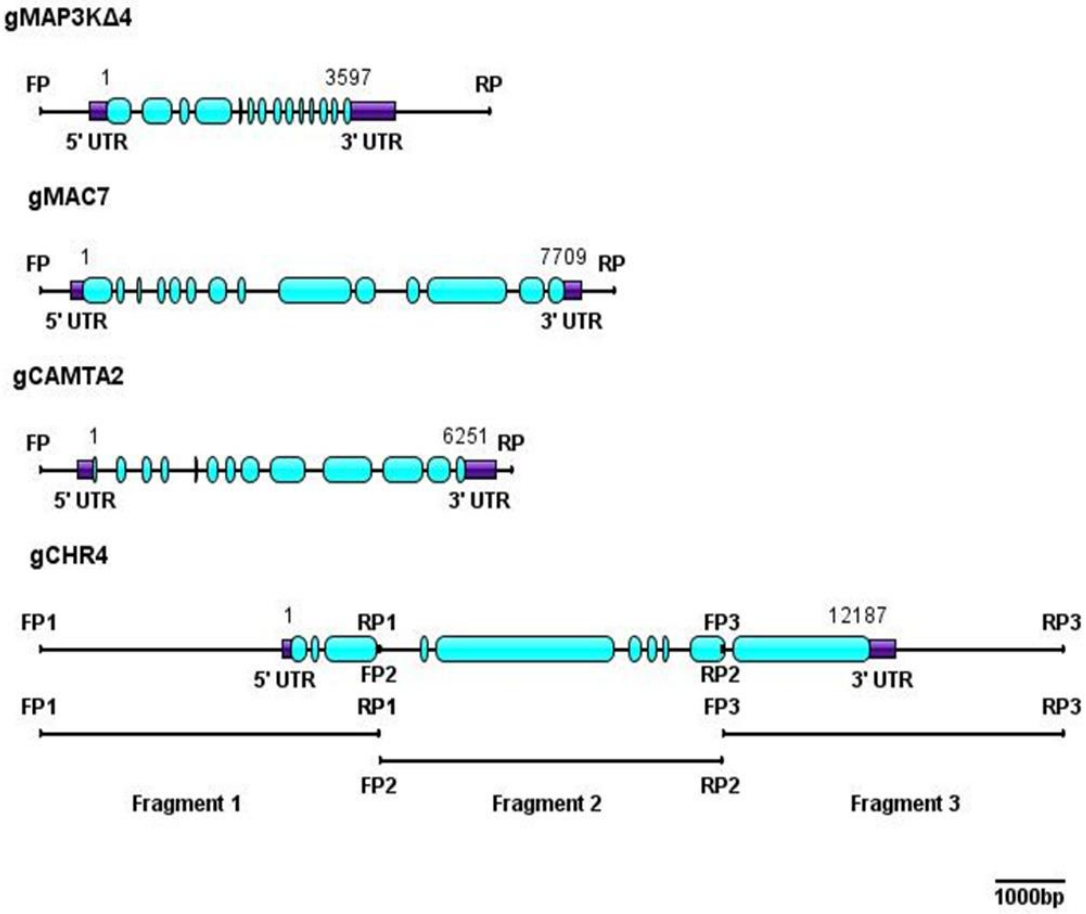


Figure 5.8: Cloning of EMS mutant candidate genes
Cloning schematic of the candidate genes *MAP3KΔ4* (*AT4G23050*), *MAC7* (*AT2G38770*), *CAMTA2* (*AT5G64220*) and *CHR4* (*AT5G44800*) responsible for the susceptibility phenotype of mutants identified in an *Albugo candida* isolate AcEM2 screen of *Arabidopsis thaliana* Col-*eds1.2* plants. Genes were cloned from the forward primer (FP), located 1496, 640, 781 and 3683 bp upstream of the start codon respectively to the reverse primer (RP) located 1671, 693, 698 and 2862 bp downstream of the stop codons. *CHR4* was cloned in three overlapping fragments which were fused to make a full-length construct. Genes are numbered from the start codon- stop codon, light blue boxes represent exons, purple boxes represent the 5' and 3' untranslated regions, scale bar represents 1000 base pairs.

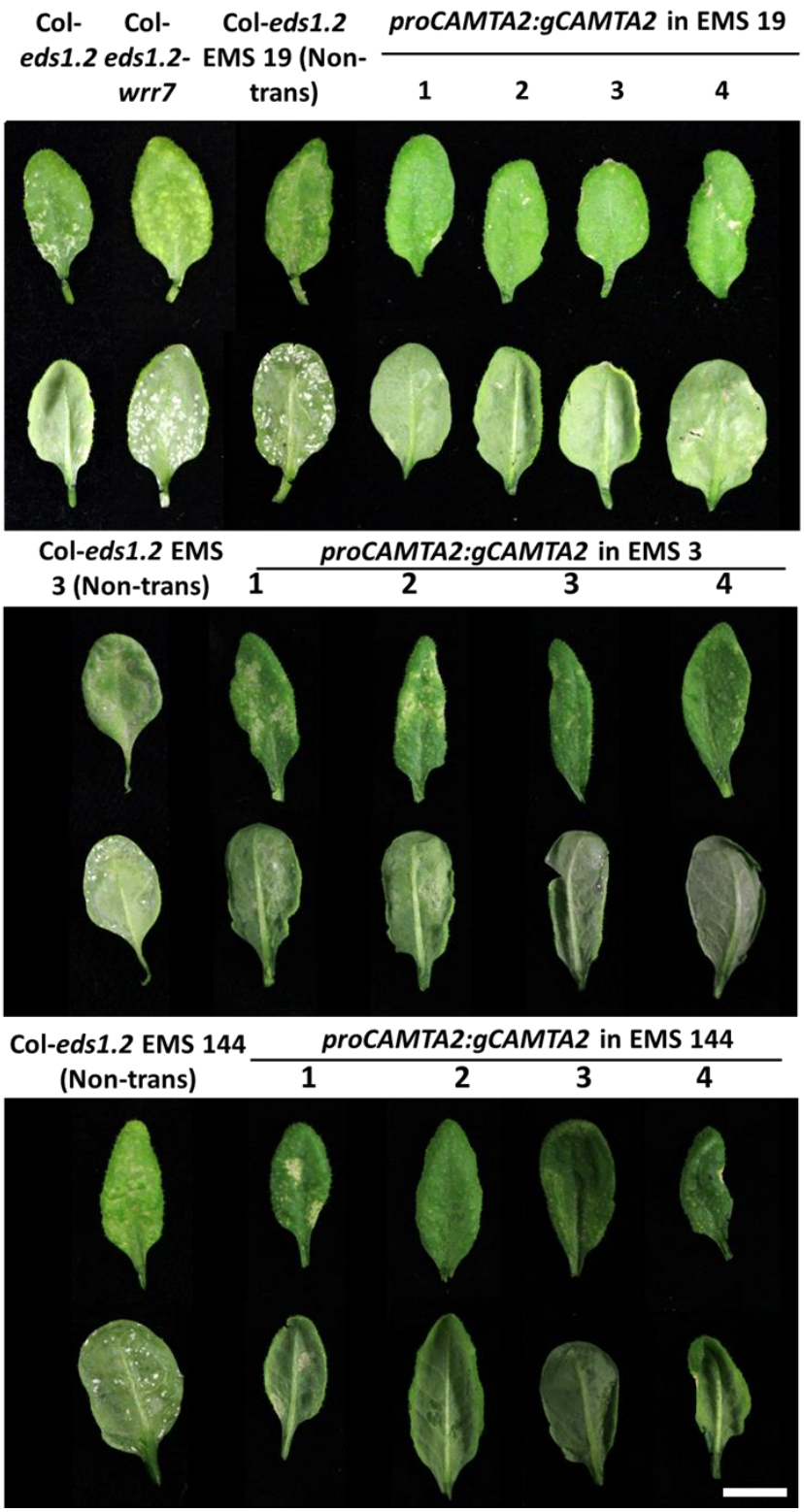


Figure 5.9: Complementation of susceptible EMS mutant lines with genomic *CAMTA2*

Leaf phenotype images of four representative independent T_1 *Arabidopsis thaliana* Col-*eds1.2* EMS *CAMTA2* mutant lines complimented with *gCAMTA2* following infection by *Albugo candida* isolate AcEM2, 10 days post infection. *gCAMTA2* was cloned with a promoter region expending 781 bp upstream of the start codon and extending 698 bp downstream of the stop codon. Non-transgenic EMS mutant lines were selected from the T_1 seed set that transgenic seeds were selected from. Scale bar represents 1cm.

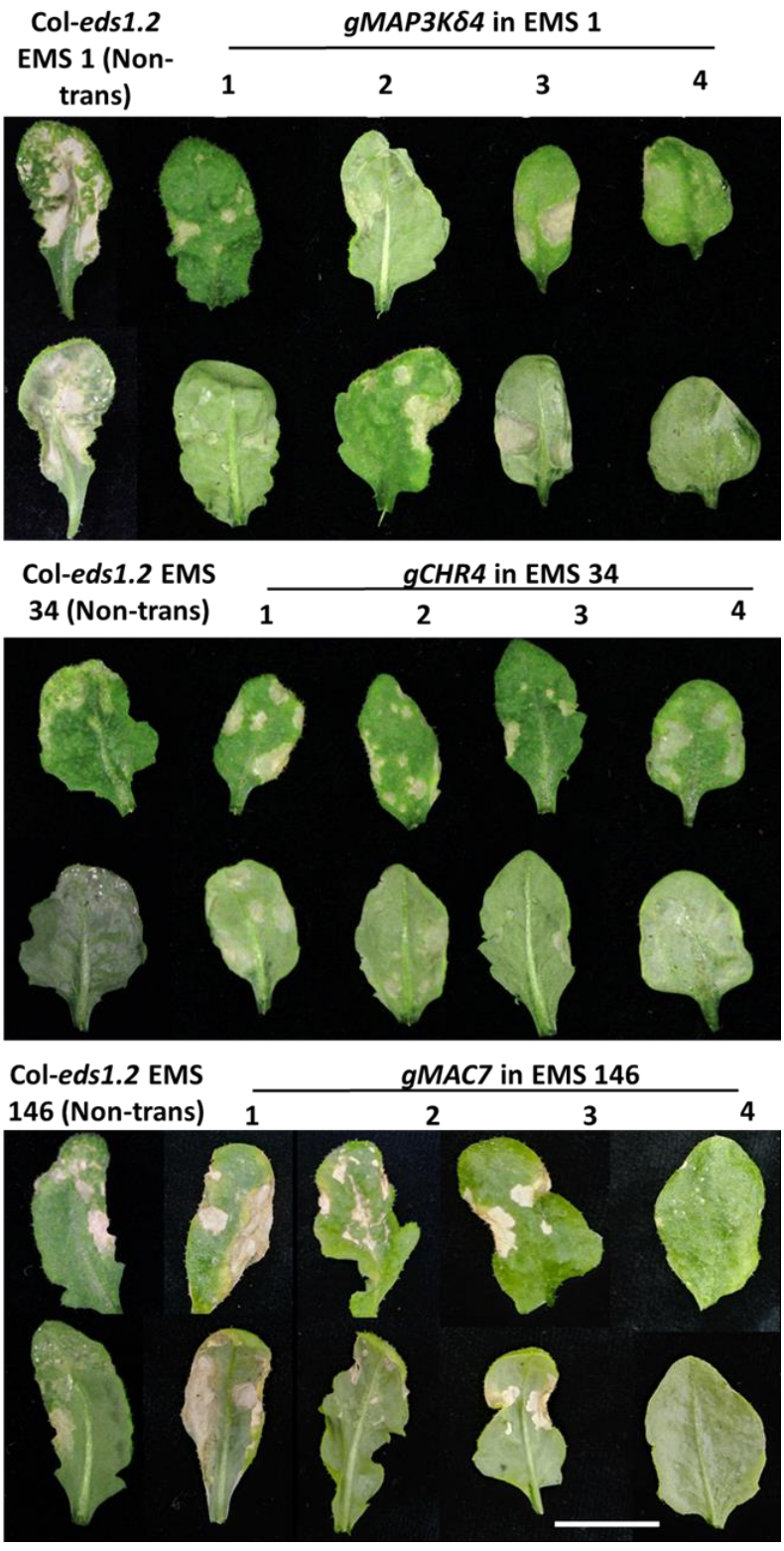


Figure 5.10: Complementation of susceptible EMS mutant lines with genomic *MAP3Kδ4, CHR4* and *MAC7*

Leaf phenotype images of four representative independent T_1 *Arabidopsis thaliana Col-eds1.2* EMS mutant lines complimented with *gMAP3Kδ4, gCHR4 and gMAC7* cloned 1496, 3683 and 640 bp upstream of their start codons to 1671, 2862 and 693 bp downstream of their respective stop codons. Images were taken following infection by *Albugo candida* isolate AcEM2, 10 days post infection. Non-transgenic EMS mutant lines were selected from the T_1 seed set that transgenic seeds were selected from. Scale bar represents 1cm.

Table 5.3: Complementation of Col-*eds1.2* EMS mutant lines with their candidate genes

Table showing the number of T_1 Col-*eds1.2* EMS mutant lines transformed with their identified candidate mutant gene. T_1 plants were then screened with *Albugo candida* isolate AcEM2 and the number of fully resistant seedlings was recorded.

Original line	Construct transformed with	Total number of T_1 lines recovered	Number of resistant T_1 lines
Col-eds1.2 EMS 1	*gMAP3Kδ4*	4	4
Col-eds1.2 EMS 3	*gCAMTA2*	15	13
Col-eds1.2 EMS 19	*gCAMTA2*	36	25
Col-eds1.2 EMS 34	*gCHR4*	7	5
Col-eds1.2 EMS 144	*gCAMTA2*	13	10
Col-eds1.2 EMS 146	*gMAC7*	7	7

CAMTA2 but not *CAMTA1* or *CAMTA3* are required for *WRR7* immunity

To further confirm our findings that these genes are involved in the *WRR7* immune signalling pathway, we obtained T-DNA knock out lines of each of the candidate genes, except MAC7 which is seedling lethal when knocked out. Interestingly, *CAMTA2* has been reported to have functional redundancy with *CAMTA1* and *CAMTA3* (Kim et al., 2017). Therefore, we obtained T-DNA KO lines of *CAMTA1* (SALK_008187) and *CAMTA3* (SALK_001152) as well as that of *CAMTA2* (SALK_007027). Each T-DNA KO line was crossed with CW14, the recombinant inbred line derived from Col-5 x Ws-2, that contained only the Col-5 allele of *WRR7* but not the loci containing *WRR4A* or *WRR5A* and *WRR5B*. F_1 crosses were then selfed to generate F_2 populations that were segregating for the T-DNA KO insertion. F_2 populations were subsequently screened with AcEM2. If the candidate gene is required for *WRR7* mediated immunity, we would expect a homozygous F_2 line for the T-DNA insertion to impose susceptibility in a crossed population derived from a CW14 x T-DNA KO line of interest cross at a ratio of 1:64 as there are three segregating loci of interest (*WRR4, WRR5* and *CAMTA2*).

Interestingly, *CAMTA2* is believed to be functionally redundant with *CAMTA1* and *CAMTA3.* However, neither of these two genes were mutated in any of EMS 19, 3 or 144 (Table 5.2) suggesting that neither of these transcription factors were involved in the *WRR7* immune response. To test this finding, we crossed T-DNA KO mutants of *CAMTA1, CAMTA2* and *CAMTA3* with CW14. We found that only F_2 populations derived from the CW14 x Col-*camta2* cross contained susceptible plants (Table 5.4). No susceptible F_2 plants were identified in CW14 crossed with either *camta1* or *camta3* T-DNA KO lines (Table 5.4). This finding highlights the fact that *CAMTA2* alone is required for *WRR7* immune signalling and does not act in a functionally redundant manner with *CAMTA1* or *CAMTA3* when regulating the *WRR7* immune response.

Table 5.4: Identification of susceptible seedlings in F_2 Populations derived from CW14 crossed with T-DNA knock out lines of *camta1/2/3*

Table showing the estimated number of F_2 seedlings screened for susceptibility to *Albugo candida* isolate AcEM2 and the number of susceptible seedlings identified. F_2 populations were derived from an original cross between CW14 and T-DNA knock out lines of *camta1/2/3* (SALK_008187, SALK_007027 and SALK_001152).

F2 line	Estimated number of plants screened	Number of susceptible plants identified	Glabrous segregating
CW 14 x Col-*camta1*	>6000	0	Yes
CW 14 x Col-*camta2*	>6000	5	Yes
CW 14 x Col-*camta3*	>6000	0	Yes

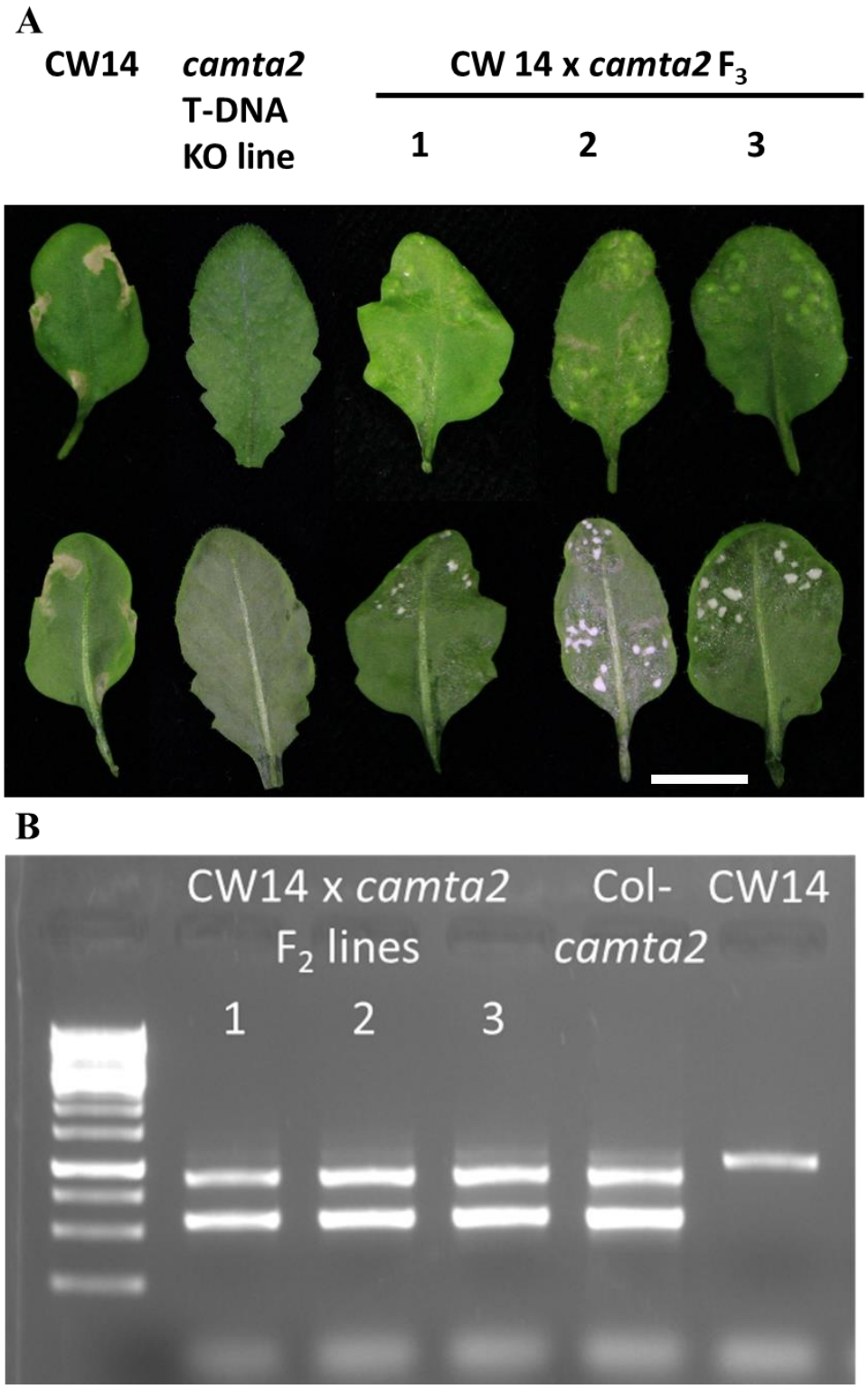


Figure 5.11: *CAMTA2* is required for *WRR7* immunity in *Arabidopsis thaliana*
A - Phenotypic images of susceptible F_3 plants identified that were derived from an original cross between CW14 and Col-*camta2* (SALK_007027). Images were taken 10 days post inoculation with *Albugo candida* isolate AcEM2. Scale bar represents 1 cm. B- Gel electrophoresis image of *camta2* genotyping results from DNA extracted from susceptible F_3 seedlings and their progenitor lines. A 1 kb ladder was used.

Expression analysis reveals that several EMS mutants have impaired *WRR7* expression

Several of the genes we identified in the *Col-eds1.2* EMS mutagenesis screen are transcriptional regulators and are involved in the *WRR7* immune mechanism (Table 5.2). These include Ca^{2+} responsive transcription factor CAMTA2, Chromatin remodelling protein, CHR4 and spliceosome associated protein MAC7. Therefore, to determine whether the susceptible phenotypes in their corresponding EMS mutant populations was due to the mis-regulation of *WRR7* at the transcriptional level, we performed a RT-qPCR analysis on these mutant lines. *WRR7* expression was measured 3- and 5-days post infection with *A. candida* isolate AcEM2 and compared to the *WRR7* transcript levels of resistant *Col-eds1.2* plants.

We found that, in comparison to Col-*eds1.2* plants, the *camta2* mutant lines EMS 19 and EMS 144 did not show up-regulation of *WRR7* after infection with AcEM2 at 3 or 5 dpi (Fig 5.12). We also observed that *WRR7* did not show significant up-regulation in the *chr4* and *mac7* mutant lines EMS34 and EMS146 3dpi with AcEM2 (Fig 5.12). However, these lines did show significant up-regulation of *WRR7* transcripts 5 dpi with *A. candida* (Fig 5.12). This is likely due to the increased number of infected cells because of the weaker immune response. A similar response was observed in the *wrr7* mutant line EMS 138 (Fig 5.12), which was unexpected as the S8F mutation was believed to be structural in its nature and not effecting the expression of *WRR7*. The *mak3kδ4* mutant, EMS 1, showed only minor levels of *WRR7* transcript reduction compared to *Col-eds1.2* 3dpi and significant up-regulation 5 dpi (Fig 5.12). This was also likely due to the increased colonisation of the pathogen in this mutant line. Therefore, MAP3kδ4 is likely active downstream of WRR7 during *WRR7* mediated immunity.

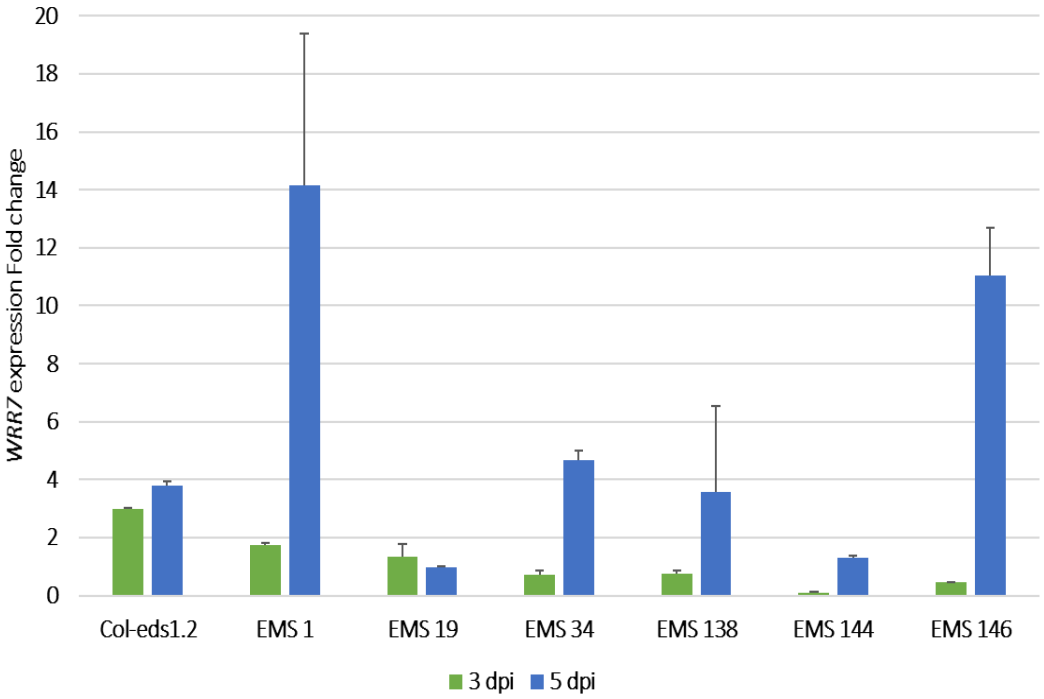


Figure 5.12: Expression analysis of EMS mutant lines reveals reduced expression of *WRR7* in *camta2, chr4* and *mac7* mutants
WRR7 expression fold change in Col-*eds1.2* mutant lines following infection by *Albugo candida* isolate AcEM2 3- and 5-days post infection (dpi). *WRR7* expression was analysed using the ΔΔCT method and expression was normalised against housekeeping gene *PP2AA3* and expression levels of AcEM2 infected samples were compared against the expression levels of mock inoculated plants of the same age. Error bars represent standard error.

CAMTA2 interacts with CaM2/3

We have shown that CAMTA2 regulates the expression of *WRR7* following infection by *A. candida* isolate AcEM2. This novel finding reveals that decoding of specific Ca^{2+} signatures during *A. candida* infection by CAMTA2 results in activation of *WRR7*. The link between Ca^{2+} signalling and the resulting CAMTA2 driven activation of *WRR7* is therefore a missing part of this signalling pathway. CAMTA transcription factors interact with Calmodulin proteins that can determine changes in Ca^{2+} homeostasis (Galon, Snir and Fromm, 2010). Therefore, we hypothesised that CAMTA2 is interacting with one of the calmodulin proteins after *A. candida* infection which stimulates its activation leading to the upregulation of *WRR7*. To test this hypothesis, we determined whether there was any differential expression of calmodulin (*CaM*) genes in *Col-eds1.2-wrr7* plants, 2, 4, and 6 days post infection using the RNA sequence dataset described earlier. We determined that, of the seven *CaM* genes present in the *A. thaliana* Col-0 genome only *CaM2* showed differential expression following *A. candida* infection (Table 5.5). Intriguingly, this gene fell into cluster H, which also shows significant enrichment for regulation by CAMTA transcription factors (Fig 5.4, Table 5.1).

Therefore, to determine whether CaM2 associates with CAMTA2 *in planta* we cloned, and epitope tagged *CaM2*, and closely related *CaM3* as well as *CAMTA2* and *CAMTA3* (Fig 5.13). We then tested whether these proteins interact using Co-immunoprecipitation experiments after transient co-expression in *N. benthamiana*. We found that CaM2-V5 and CaM3-V5 both immunoprecipitated with CAMTA2-HF as well as CAMTA3-HF using FLAG beads but the YFP-V5 control could not IP with either CAMTA2-HF or CAMTA3-HF (Fig 5.14). Therefore, we hypothesise that after infection by *A. candida* CaM2/3 is able to bind to Ca^{2+} following a change in Ca^{2+} homeostasis and subsequently interacts with CAMTA2, stimulating the upregulation of *WRR7* that leads to defence activation.

Table 5.5: Calmodulin 2 is the only calmodulin gene to change in expression in *Arabidopsis thaliana Col-eds1.2-wrr7* plants after infection by *Albugo candida* isolate AcEM2

Expression fold change values (log2) from RNA sequencing data of Calmodulin genes (*CaM*) in *Arabidopsis thaliana Col-eds1.2-wrr7* 2, 4- and 6-days post infection (dpi) with *Albugo candida* isolate AcEM2. Differentially expressed genes were determined to have a log2 fold change value >1 or <-1.

Calmodulin gene	ATG number	Differentially expressed gene	Fold change (Log2)		
			2dpi	4dpi	6dpi
CaM1	*AT5G37780*	No	0.010	0.048	-0.142
CaM2	*AT2G41110*	Yes	1.764	2.384	1.500
CaM3	*AT3G56800*	No	0.224	0.495	0.204
CaM4	*AT1G66410*	No	0.199	0.231	0.009
CaM5	*AT2G27030*	No	-0.062	-0.152	-0.166
CaM6	*AT5G21274*	No	0.029	0.538	0.166
CaM7	*AT3G43810*	No	0.145	0.100	-0.605

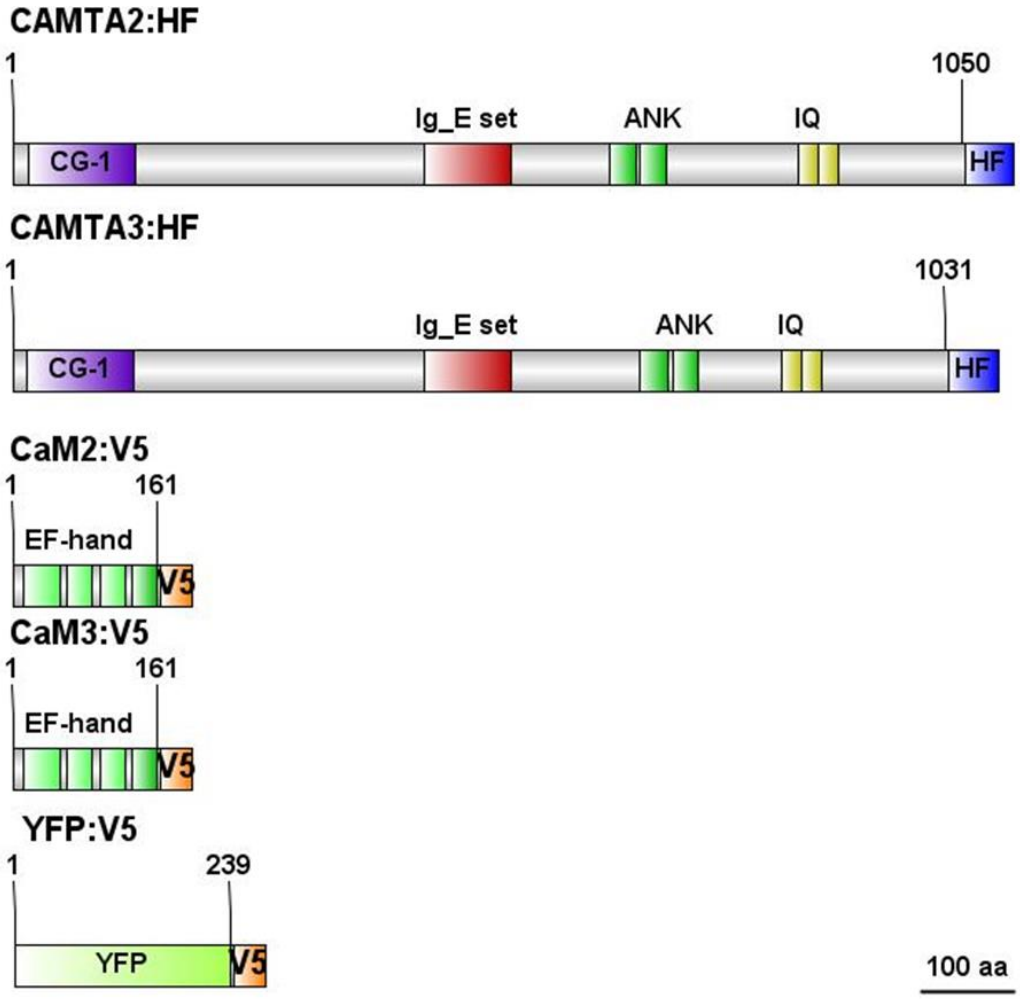


Figure 5.13: Cloning and epitope tagging of CAMTA transcription factors and Calmodulin proteins

Cloning schematic of CAMTA and Calmodulin (CaM) proteins cloned into *Agrobacterium tumefaciens* strain GV3101. Proteins domains were predicted using SMART software (Letunic, Ivica, Doerks and Bork, 2015). Scale bar represents 100 amino acids.

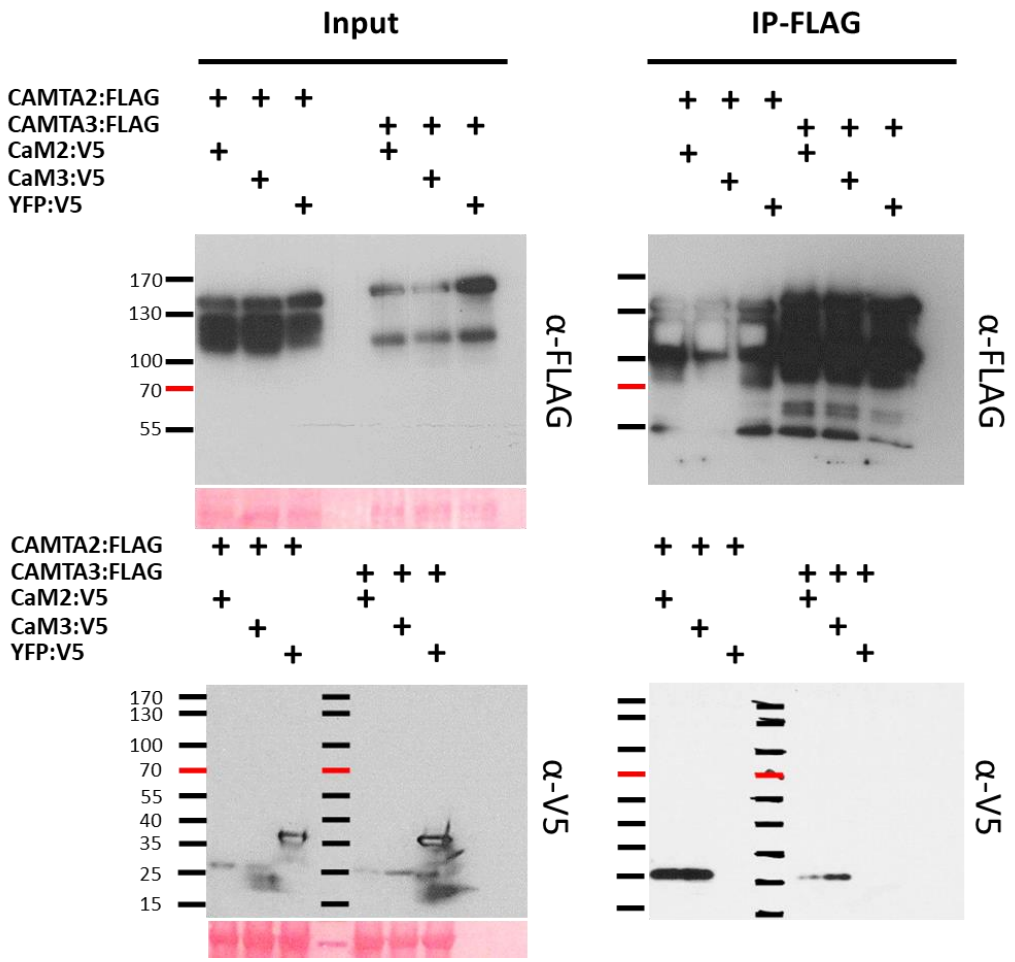


Figure 5.14: CAMTA2 and CAMTA3 associate with Calmodulin 2 and Calmodulin 3

Immunoblots of FLAG co-immunoprecipitation experiments of transiently overexpressed epitope Histidine-FLAG tagged CAMTA transcription factors with V5 epitope tagged Calmodulins or yellow fluorescent protein (YFP) in *Nicotiana benthamiana*. Ladder represents kDa and ponceau S stained membranes are shown below input blots.

ADR1s and NRGs are not required for *WRR7* mediated immune signalling

The *NRG* and *ADR1* families are required for the downstream signalling of TNLs but not CNLs (Castel et al., 2018). To test whether either the *NRG* family or *ADR1* gene family are required for *WRR7* signaling, the recombinant inbred line containing just the *WRR7* loci CW14 was crossed with triple knock out lines of the *NRG* and *ADR1* genes. F2 populations derived from this cross were then screened to identify any susceptible individuals. In the F2 populations we would expect a susceptible: resistant ratio of plants of 1:64 for the CW 14 x *nrg triple* cross and a 1:1024 ratio for the CW14 x *adr1 triple* cross if they are involved in *WRR7* signaling due to the segregation of *nrgs/adr1s* from the CW RILs and the segregation of the two WRR loci absent from the CW lines but present in the Col-*nrg/adr1 triple* mutant backgrounds. No susceptible plants were identified in either F_2 populations therefore *NRG*s and *ADR1s* are not required for WRR7 downstream immune responses (Table 5.6).

Table 5.6: ADR1 and NRG proteins are not required for *WRR7* mediated immunity

Table of resistant: susceptible phenotypes observed in F_2 populations of crosses derived from *WRR7* single loci containing recombinant inbred line CW14 (from original Col-5 x Ws-2 cross) crossed with *triple nrg* (*nrg1.1*, *nrg1.2* and *nrg1.3*) or *adr1 triple* (*adr1*, *adr1-L1* and *adr1-L2* mutant lines).

F2 line	Estimated number of plants screened	Number of susceptible plants identified	Glabrous 3:1 segregation ratio
CW14 x *nrg triple*	6000	0	Yes
CW14 x *adr1 triple*	10,500	0	Yes

The C442Y mutation in WRR7 induces autoimmunity

WRR7 contains an integrated LIM-Peptidase that is similar to the LIM-Peptidase integrated domain that is present in WRR5B. WRR5B has an autoimmune allele *chs3*-2D which causes autoactivation leading to an autoimmune response (Fig 3.7). The *chs3*-2D mutation is a non-synonymous G to A nucleotide substitution causing a single amino acid substitution of a highly conserved cysteine to tyrosine at amino acid position 1340, between the LIM and Peptidase domains of WRR5B (Bi, D. et al., 2011). This site is conserved between WRR5B and WRR7 and the WRR5B C1340 corresponds to C442 in WRR7 (Fig 5.15). Therefore, we hypothesised that the induction of a C442Y mutation in WRR7 would result in an autoimmune response. To test this hypothesis, we generated a WRR7 C442Y mutant and a WRR7 construct C-terminally tagged with HF and transiently expressed these constructs in *N. benthamiana* and *N. benthamiana-nrg1*.

We found that WRR7-C442Y mutants caused an autoimmune response in *N. benthamiana* and surprisingly C-terminal tagging of WRR7 with a HF epitope tag resulted in autoimmunity but unaltered *WRR7* did not induce autoimmunity (Fig 5.16). We also confirmed that these autoimmune responses could not be abolished in *N. benthamiana-nrg1* lines compared to the autoimmunity induced by the co-infiltration of WRR5B-*chs3*-2D with WRR5A which was inhibited in *N. benthamiana-nrg1* plants (Fig 5.16), confirming our finding that NRG1 is not required for the downstream functioning of WRR7.

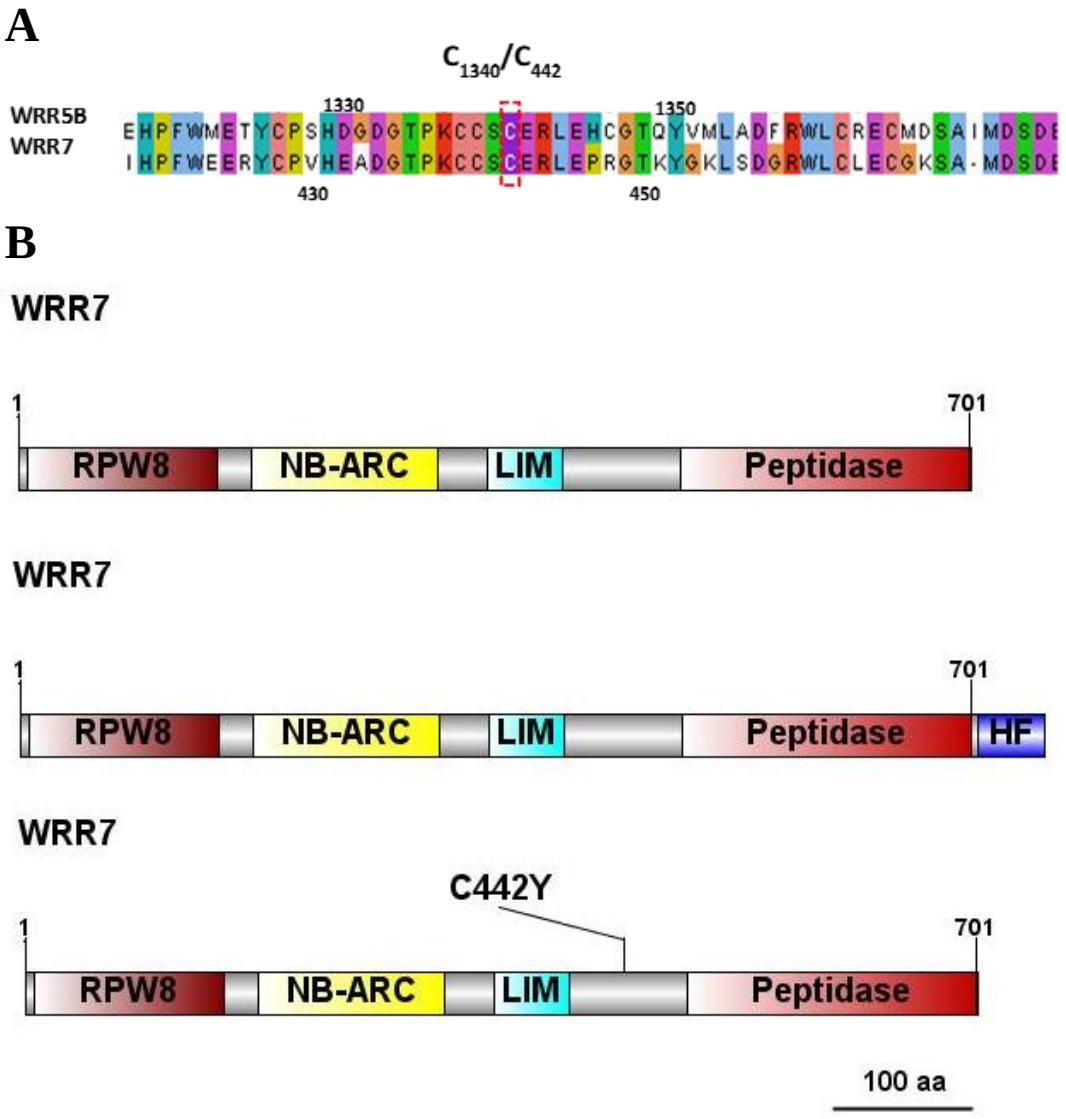
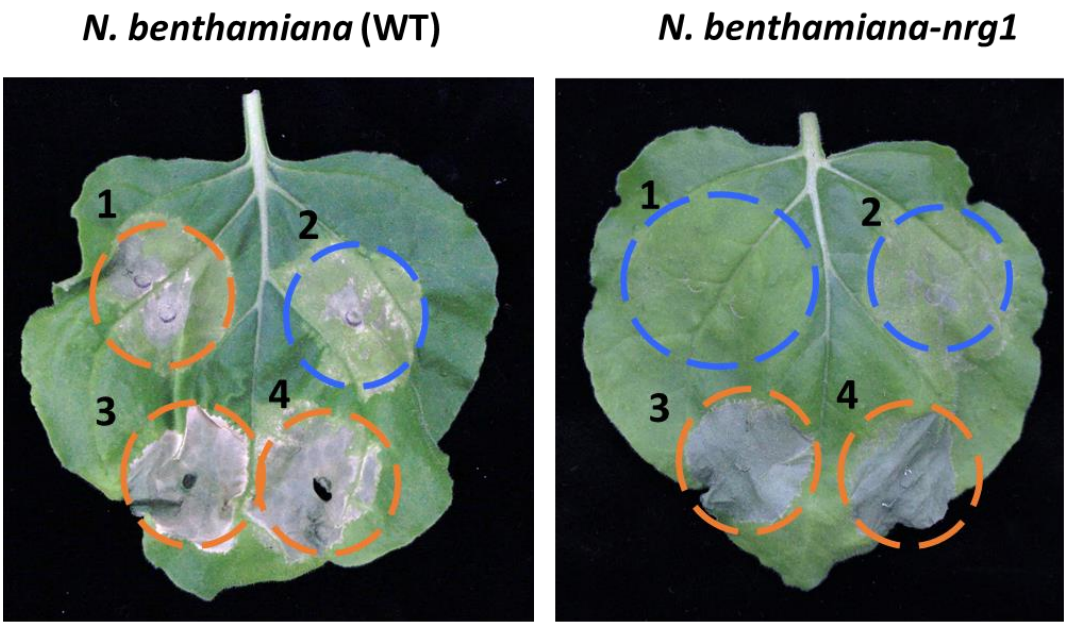


Figure 5.15: Cloning of *WRR7*
A - muscle alignment of WRR5B and WRR7 full length amino acid sequences at the site of the WRR5B-C1340Y mutation. Colours of amino acid residues are based on the Clustal X colour scheme. B - Protein schematic of WRR7 constructs cloned into *Agrobacterium tumefaciens* strain GV3101. Proteins domains were predicted using SMART software (Letunic, Ivica, Doerks and Bork, 2015). Scale bar represents 100 amino acids.



Infiltration number	Constructs infiltrated	Response in *N. benthamiana* (WT)	Response in N. *benthamiana-nrg1*
1	WRR5A and WRR5B-*C1340Y*		
2	WRR7		
3	WRR7-HF		
4	WRR7-C442Y		

Figure 5.16: The C442Y mutation in WRR7 induces autoimmunity and NRG1 is not required for *WRR7* mediated immunity
Images of cell death assays following transient expression of *Agrobacterium tumefaciens* strain GV3101 containing the constructs of interest (table). Images were taken three days post infiltration with *A. tumefaciens*. Blue circles and corresponding cells in the table represent areas with no observable cell death following agro-infiltration, orange circles and cells show leaf areas showing a cell death response.

The EMS 138 S8F mutation abolishes autoimmunity caused by the WRR7-C442Y mutation

The S8F mutation identified in EMS 138 is located in the first α-helix of the RPW8 domain in WRR7. The first α-helix of the CNL protein, ZAR1 has been shown to form a barrel-like pore structure capable of integrating into membranes after the formation of the resistosome (Wang, J., Hu, et al., 2019). Earlier we showed that the C442Y mutation causes autoimmunity in WRR7 (Fig 5.16). To test whether the first α-helix of WRR7 behaves in a similar manner to the ZAR1 α-helix, we cloned and epitope tagged WRR7 with YFP. We then determined the cellular location of WRR7 by confocal microscopy. In addition, we generated mutant versions of YFP epitope tagged WRR7, containing the S8F mutation and the C442Y mutation (Fig 5.17), to determine whether the S8F mutation was capable of abolishing autoimmunity induced by the C442Y mutation.

Here we show that N-terminally and C-terminally YFP tagged WRR7 is localised to the plasma membrane as well as the nucleus and that both WRR7-S8F and WRR7-S8F-C442Y constructs showed the same localisation pattern (Fig 5.18). Intriguingly, transient expression of C-terminally tagged WRR7 with YFP also caused an autoimmune response that was not observed in WRR7 constructs N-Terminally tagged with YFP (Fig 5.19). We also confirmed that the WRR7-S8F mutation was able to abolish the autoimmunity induced by the WRR7-C442Y mutation as well as autoimmunity triggered by C-terminal tagging of WRR7 (Fig 5.19).

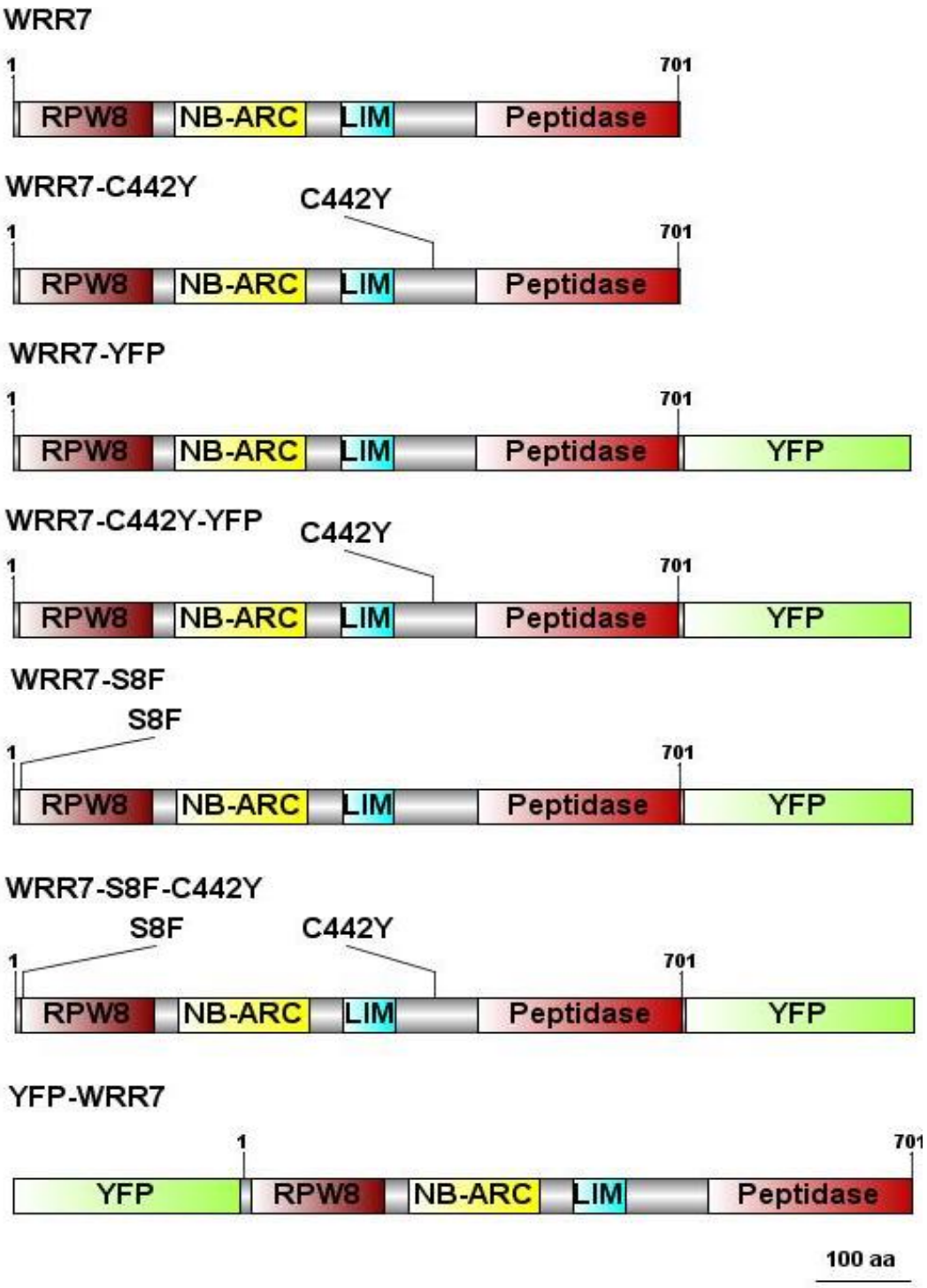


Figure 5.17: Cloning and epitope tagging of WRR7 mutants
Protein schematic of WRR7 mutations and epitope tags transformed into *Agrobacterium tumefaciens* strain GV3101. Proteins domains were predicted using SMART software (Letunic, Ivica, Doerks and Bork, 2015). Scale bar represents 100 amino acids.

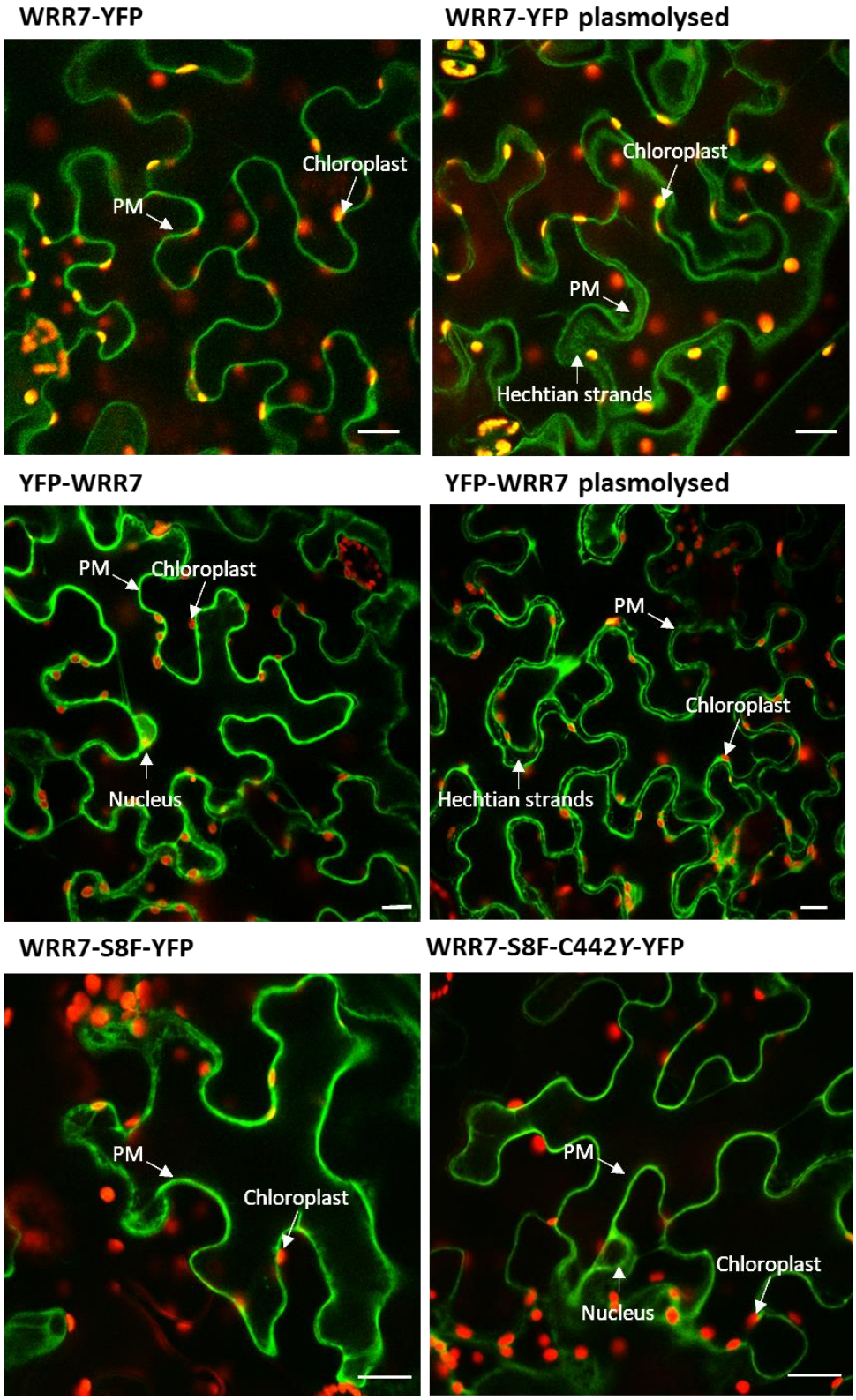
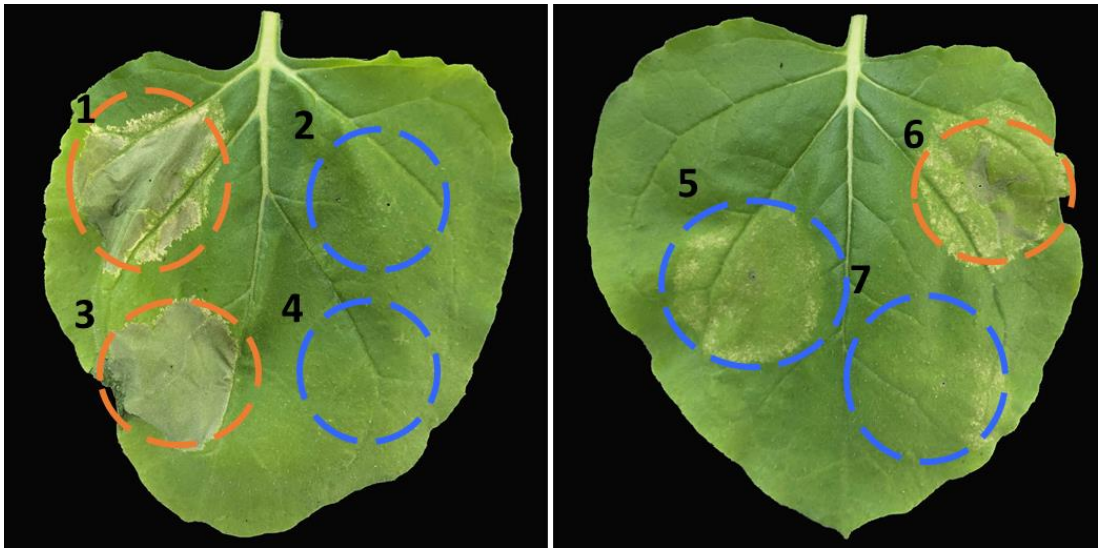


Figure 5.18: WRR7 localises to the plasma membrane
Confocal microscopy images of transiently overexpressed YFP epitope tagged constructs in *Nicotiana benthamiana* leaf epidermal cells, 2 days post infiltration with *Agrobacterium tumefaciens* strain, GV3101 containing WRR7 variants. Red shows chloroplast autofluorescence and green shows YFP localisation. Examples of the plasma membrane (PM), nucleus, chloroplast and hechtian strands are labelled and the scale bars represents 20µm.



Infiltration number	Constructs infiltrated	Response in *N. benthamiana* (WT)
1	WRR7-YFP	
2	WRR7-S8F-YFP	
3	WRR7-C442Y-YFP	
4	WRR7-S8F-C442Y-YFP	
5	WRR7	
6	WRR7-C442Y	
7	YFP-WRR7	

Figure 5.19: The S8F mutation in WRR7 inhibits the cell death induced by the WRR7 autoimmune mutant WRR7-C442Y
Cell death assays of transiently expressed *Agrobacterium tumefaciens* strain GV101 containing WRR7 YFP epitope tagged constructs (see table) in *Nicotiana benthamiana* leaves 3 days post inoculation. Orange represents assays showing a cell death response, blue represents infiltrated areas showing no cell death response.

Discussion

In chapter 3 (Fig 3.10), we identified *WRR7* as being responsible for the necrotic resistance phenotype observed in recombinant inbred lines derived from the Col-5 x Ws-2 cross. In this chapter, we used a forward genetic screen to identify other potential proteins that are involved in the *WRR7* resistance mechanism (Fig 5.6). We identified several proteins that are involved in the WRR7 resistance pathway (Table 5.2). These include proteins associated with the transcriptional regulation, including Ca^{2+} responsive transcription factor CAMTA2, a chromatin remodelling protein (CHR4) and a protein associated with the spliceosome machinery MAC7. As well as proteins involved in transcript regulation, we also identified a mitogen-activated protein kinase kinase kinsase (MAP3Kδ4) that is important in regulating ABA signalling.

The Ca^{2+} responsive transcription factor CAMTA2 was identified in three *Col-eds1.2* mutant lines (Table 5.2). There are six *CAMTA* genes in the *A. thaliana* Col-0 genome, all six are predominantly localised to the nucleus and act as transcription factors (Bouché et al., 2002). CAMTA proteins contain a string of conserved domains including nuclear localisation signal (CG-1), DNA binding domain, TIG domain (PDB), Ankyrin repeats and C-terminal IQ domains (Finkler, Ashery-Padan and Fromm, 2007; Poovaiah et al., 2013). These domains are each important to the functioning of the proteins. The CG-1 nuclear localisation signal is not only important for the correct cellular trafficking of the protein to the nucleus but is also important in regulating the binding of the CAMTA proteins to the conserved CG(C/T)CG promoter elements of target genes (Finkler, Ashery-Padan and Fromm, 2007). The Ankyrin repeats are involved in protein-protein interactions and the C-terminal IQ domains regulate the binding of CAMTAs with calcium decoding proteins such as calmodulin (CaM) and calmodulin like (CaM-like) proteins (Finkler, Ashery-Padan and Fromm, 2007; Poovaiah et al., 2013). The 6 CAMTA genes present in *A. thaliana* can be split into two subgroups based on the presence or absence of the TIG domain, which is absent in CAMTA1/2/3 but present in CAMTA4/5/6 (Kim et al., 2017). CAMTA2 is part of the first subgroup which lacks a TIG domain. The two other CAMTA proteins in this subgroup, CAMTA1 and CAMTA3 have been shown to act in a redundant manner with CAMTA2 during both biotic and abiotic stress responses (Du, Liqun et al., 2009; Tokizawa et al.,

2015; Kidokoro et al., 2017; Kim et al., 2017). Contrary to previous studies, we identified CAMTA2 as being solely responsible for the activation of *WRR7* in *A. thaliana* plants following infection by *A. candida* and not acting in a functionally redundant manner with CAMTA1 and CAMTA3 (Table 5.4). Therefore, we propose that CAMTA2 activates *WRR7* expression following *A. candida* infection by binding to the CGCGT motif that lies 93 bp upstream of the start codon in the *WRR7* promoter region (Fig 5.5). The most studied CAMTA transcription factor in plants is CAMTA3, which has been shown to negatively regulate immunity associated genes and activate developmental genes (Yuan, Du and Poovaiah, 2018).

The fact that a CAMTA transcription factor is involved in regulating *WRR7* expression following infection by *A. candida* strongly suggests that *WRR7* is induced proceeding an *A. candida* induced change in Ca^{2+} homeostasis. CAMTA transcription factors bind to Calmodulin (CaM) proteins that sense the presence of Ca^{2+} by sequestering it in EF-hand domains (Rahman et al., 2016; Kudla et al., 2018). Calmodulin (CaM) proteins form a small family of signalling proteins of which there are 7 in Arabidopsis that form 4 isoforms: CaM1/4, CaM2/3/5, CaM6 and CaM7 (Poovaiah et al., 2013; La Verde, Dominici and Astegno, 2018). Therefore, it is likely that an alteration to the cellular Ca^{2+} concentrations caused during *A. candida* infection is sensed by CaM proteins that then bind to CAMTA2, stimulating the activation of *WRR7*. We identified *CaM2* as the only *CaM* gene to have increased expression after *A. candida* infection (Table 5.5) and showed that both CaM2 and CaM3 proteins can physically associate with CAMTA2 *in planta* (Fig 5.14). Suggesting that the CaM2/3/5 isoform can sense *A. candida* through changes in Ca^{2+} homeostasis, this protein then binds to CAMTA2 causing *WRR7* activation.

As well as CAMTA2, we identified a chromatin remodelling complex, CHR4 as being involved in the regulation of *WRR7* transcripts (Table 5.2). CHR4 is part of CHD subfamily II chromatin remodelling complexes that have been shown to bind to H3K27me3 marks and activate genes, in an antagonistic manner to Polycomb group proteins (Hennig and Derkacheva, 2009; Zhang, H. et al., 2012; Carter et al., 2018; Jing, Lin and Guo, 2019). The closest homolog of CHR4 in Rice (OsCHR729) has also been shown to regulate H3K27me3 trimethylation (Hu et al., 2012; Hu, Lai and Zhu, 2014). A recent study into the *chr4-2* T-DNA mutant line revealed that

when *chr4* is mutated the *wrr7* loci is bivalently hypermethylated for both H3K27me3 and H3K4me3 marks (Sang et al., 2020). Therefore, it is likely that CHR4 acts in an antagonistic action to the repression of H3K27me3 methylated genes. We propose that after *A. candida* infection, CHR4 binds to H3K27me3 sites in the *WRR7* promoter allowing CAMTA2 to bind to the CGCGT motif, stimulating the expression of *WRR7* due to its primed H3K4me3 state. How CHR4 recognises the presence of *A. candida* is yet to be elucidated although a scenario where CHR4 interacts with CAMTA2 has potential to explain how both these proteins could be responding to *A. candida* infection in *A. thaliana.*

As well as epigenetic regulation of genes, transcription can also be regulated by RNA processing. Once a pre-mRNA transcript is produced it associates with RNA binding proteins that modulate its splicing, turnover and trafficking (Staiger et al., 2013; Herzel et al., 2017). In plants it has recently been shown that both transcription and splicing occur simultaneously through co-transcriptional splicing (Zhang, Y. and Ding, 2020). This layer of RNA based control provides another layer of regulatory proteins that can control the levels of particular proteins within the cell. We identified MAC7 as a component of the *WRR7* resistance pathway (Table 5.2). MAC7 is an RNA helicase that is associated with the MOS4 spliceosome complex (Palma et al., 2007; Staiger et al., 2013; Jia et al., 2017). The mammalian homolog of MAC7, Aquarius has been extensively studied and is an integral component of the spliceosome complex (De et al., 2015). In plants, MAC7 has been shown to be important for pre-mRNA splicing as well as the biogenesis of miRNAs and *mac7-1* mutant plants showed a significant downregulation of defence associated genes (Jia et al., 2017). In addition, the MAC complex has been implicated in the recruitment of the spliceosome to the chromatin during transcription and mutants of *mac* genes result in an increase in defects in nascent RNAs (Li, S. et al., 2020; Zhang, Y. and Ding, 2020). The combination of our findings of MAC7 being important for the processing of *WRR7* transcripts as well as the finding that MOS4 is important in the regulation of *SNC1* transcripts provides increasing evidence that the MOS4 spliceosome complex plays a major role in regulating resistance gene transcripts in response to infection.

The final mutant we identified as being part of the *WRR7* resistance mechanism against *A. candida* was MAP-kinase kinase kinase MAP3Kδ4 (Table 5.2).

MAP3Kδ4 is a Raf-like MAP3K that when constitutively expressed increases plant growth (Sasayama et al., 2011). It also has an important regulatory function in ABA signalling and is known to associated with key phosphatases PP2Cs and the ABF transcription factors in the core ABA signalling pathway (Shitamichi et al., 2013; Lumba et al., 2014). As well as MAP3Kδ4, other Raf-like MAP3Ks have been shown to be core regulators of ABA signalling (Lee et al., 2014; Nguyen et al., 2019). ABA is a phytohormone that has an antagonistic association with the ETI defence hormone salicylic acid (Moeder et al., 2010). Therefore, it is not surprising that we identified a negative regulator of ABA as being involved in the downstream response of WRR7 signalling. The MAP3Kδ4 mutant EMS 1 displayed a particularly weak susceptible phenotype with the development of only a few pustules and extensive cell death (Fig 5.6). Therefore, MAP3Kδ4 is likely to only be one of several proteins that operate in repressing the ABA signalling pathway during infection. The upregulation of SA is obligate in ETI signalling; therefore, we speculate that the repression of ABA signalling by MAP3Kδ4 and other Raf-like MAP3Ks may be important not only in *WRR7* signalling but more broadly in the downstream signalling of other resistance genes. Interestingly, MAP3Kδ4 is shown to interact with four *Ralstonia psuudosolanacearum* effector proteins (Fig 5.20) in the EffectorK host-pathogen interactome dataset (González-Fuente et al., 2019). The identification that MAP3Kδ4 is important for *WRR7* signalling and is also the target of *R. psuudosolanacearum* effectors (Fig 5.20) supports the idea that MAP3Kδ4 is more generally involved in ETI immune responses and plays a key role in suppressing ABA signalling during pathogen infection. The link between an infection by *A. candida* and the subsequent suppression of ABA by MAP3Kδ4 is still unclear. However, there is potential that alterations to Ca^{2+} concentrations could be responsible for MAP3Kδ4 activity as we have already shown Ca^{2+} to be important in the *WRR7* mechanism. Alternatively, MAP3Kδ4 could sense cellular disruption by *A. candida* by sensing the presence of reactive oxygen species (ROS) released during infection because MAP3Kδ4 contains a PAS domain at its N-terminus, a domain that has previously been shown to bind ROS (Vogt and Schippers, 2015).

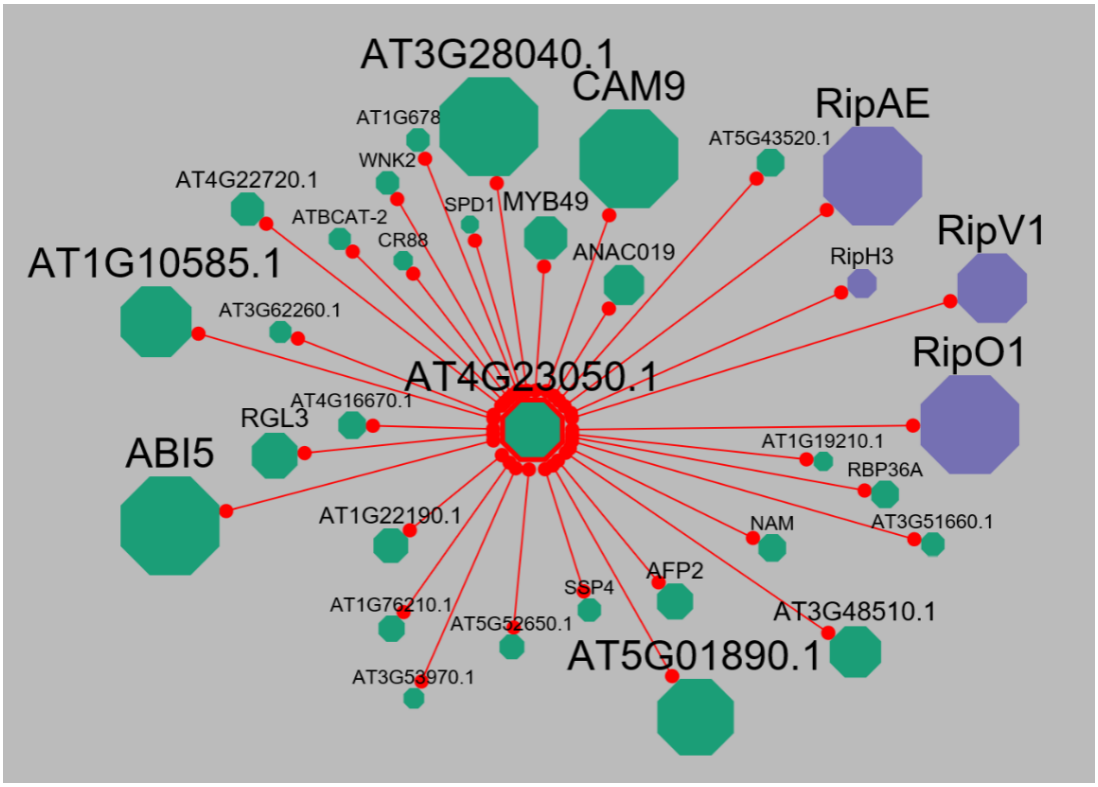


Figure 5.20: MAP3Kδ4 interacts with multiple effector proteins from *Ralstonia psuudosolanacearum*

Interactome network of MAP3Kδ4 (AT4G23050) with host (green) and pathogen effector (purple) proteins generated from the EffectorK database (González-Fuente et al., 2019). Bubble size represents the number of other interactions each protein is known to have.

We not only identified novel signalling proteins involved in the *WRR7* immune pathway, but we also identified a point mutation (S8F) in the first α-helix of the RPW8 domain (Table 5.2). The first α-helix of NLRs has recently been shown to protrude from the resistosome complex formed by ZAR1 after pentamerisation in an immune response (Wang, J., Hu, et al., 2019). Serine residues are phosphorylated during signalling events. The serine to phenylalanine point mutation that we observed in WRR7 could potentially inhibit an important phosphorylation event that is relevant to the function of the first α-helix during infection.

We have shown that the WRR7 resistance mechanism is governed by the upregulation of *WRR7* following *Albugo candida* infection (Fig 5.1). *A. candida* most likely triggers a cytosolic Ca^{2+} influx across the plasma membrane through Ca^{2+} channels such as the recently identified CNGC channels that is sensed during MTI (Meena et al., 2019; Tian, W. et al., 2019; Yu, X. et al., 2019). This cytosolic Ca^{2+} influx is then decoded by Ca^{2+} decoding proteins, most likely to be CaM2/CaM3 which interact with CAMTA2. Once this interaction takes place, CAMTA2 in combination with chromatin remodelling complex 4 (CHR4) stimulate the expression of *WRR7* (Fig 5.21). *WRR7* transcripts are then processed by the spliceosome complex that requires MAC7 for the correct processing of the transcript (Fig 5.21). Following recognition of an unknown *A. candida* effector, WRR7 then most likely homodimerizes to form a resistosome-like complex, similar to the ZAR1 resistosome complex (Wang, J., Hu, et al., 2019) that requires an intact first α-helix to integrate into the plasma membrane and enact downstream signalling. The severity of this downstream response relies on the interaction of the immune signalling network with the ABA signalling network through Mitogen-active protein kinase MAP3KΔ4 that 'ramps' the defence response making it strong enough to completely kill the invading pathogen.

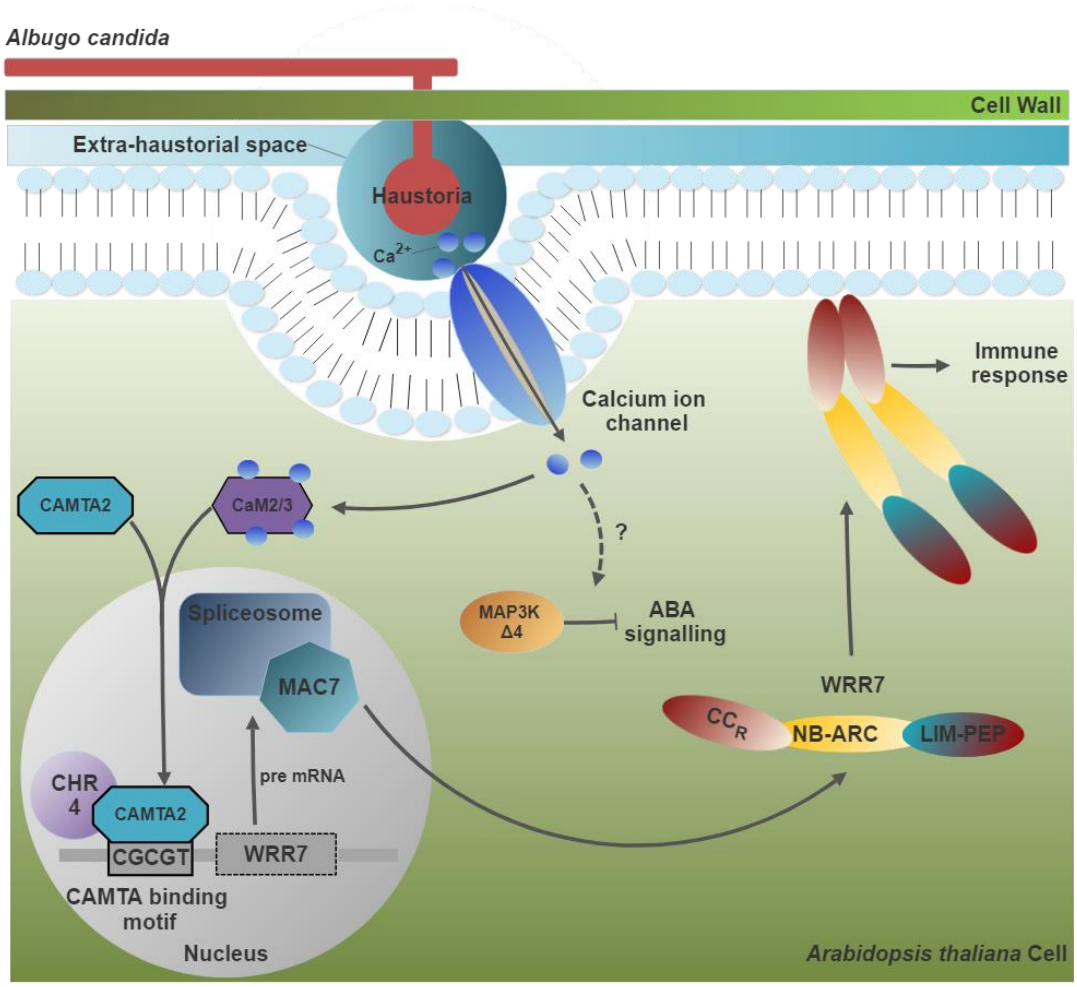


Figure 5.21: Model of *WRR7* mediated immune response to *Albugo candida*
The proposed model of *WRR7* activated immunity to *A. candida* isolate AcEM2. During infection *A. candida* causes disruption to host tissue that is sensed by a change in Ca^{2+}homeostasis, potentially via a cytosolic influx of Ca^{2+} across the plasma membrane. Ca^{2+} stimulates the repression of Abscisic acid signalling (ABA) mediated by MAP3Kδ4 and the change in Ca^{2+} concentration is decoded by Calmodulin (CaM) proteins which bind to Calmodulin binding transcriptional activator (CAMTA2). CAMTA2 the interacts with Chromatin remodelling complex 4 (CHR4) in the nucleus facilitating the binding of CAMTA2 to the CGCGT motif in the *WRR7* promoter. *WRR7* pre-mRNA is then processed by the MOS4 spliceosome complex in a MAC7 dependent manner. Mature mRNA is then translated into WRR7 protein which is translocated to the plasma membrane where it stimulates an immune response.

Chapter 6

The evolutionary history of *WRR5B* and *WRR7*

Introduction

Integrated domains (IDs) are being identified in around 8-10% of all plant NLRs and are becoming an increasingly important area of study for plant immunologists (Kroj et al., 2016; Sarris, P. F. et al., 2016; Bailey et al., 2017). IDs are enriched for domains that are the putative intracellular targets of pathogen effectors. How these particular domains are recruited into NLR architectures is not yet understood (Sarris, P. F. et al., 2016; Van de Weyer et al., 2019). Both WRR5B and WRR7 share similar LIM-Peptidase IDs and these two Resistance genes cause resistances to *A. candida* isolate AcEM2 in *A. thaliana* Col-0 plants (chapter 3, Fig 3.5 and 3.10). Therefore, we can use *WRR5B* and *WRR7* to study whether there is a common mechanism that led to the integration of the LIM-PEP domain into both genes.

In plants the LIM domain is sequentially adjoined to the Peptidase domain and forms the basis of the DA1 (meaning Big in mandarin) and DA1 related (DAR) family of proteins (Li, Yunhai et al., 2008). In *A. thaliana,* there are 8 members of this family including WRR5B (DAR4) and WRR7 (DAR5) (Srivastava and Verma, 2017). The non-Resistance gene DA1 family proteins consist of a varying number of N-terminal ubiquitin binding motifs followed by the LIM and Peptidase domains (Fig 6.1). DA1, DAR1 and DAR2 are associated with regulating organ growth including endosperm and leaf tissue due to their role in mediating endoreduplication during the cell cycle (Li, Yunhai et al., 2008; Peng, Yuancheng et al., 2015; Wang, J.-L. et al., 2017). The loss of function *da1* mutants don't display any reduction in organ size, however a dominant-negative mutant *da1-1* increases organ size (Li, Yunhai et al., 2008; Vanhaeren et al., 2017). DA1 has also been shown to interact with E3 ubiquitin ligases DA2 and Enhancer of DA1/Big Brother (EOD1/BB) and multi-monoubiquitination of the DA1 protein by these E3 ligases has been shown to activate the latent peptidase domain, subsequently cleaving both E3 ligases (Xia, T. et al., 2013; Dong, H. et al., 2017). Double *da1-1* and *eod1* mutants have a cumulative effect, exaggerating the increase in organ size observed with single *da1-1* mutants (Vanhaeren et al., 2017). The other non-resistance gene DA1 family members, DAR3, DAR6 and DAR7 have no known functions. Both *DAR4* (*WRR5B*)

and *DAR5* (*WRR7*) are resistance genes that confer resistance to *Albugo candida*. The fact that two resistance genes containing the same integrated domain are able to confer resistance against *A. candida*, suggests that this pathogen in some way interacts with one or more of the DA1 family proteins and this activity is detected via the integrated decoy domains of WRR5B and WRR7.

Decoy proteins are proteins that mimic the host targets of effectors, this allows them to bind to pathogen effector proteins and this interaction is perceived by an NLR that then stimulates a defence response (Van Der Hoorn and Kamoun, 2008). Decoy proteins often evolve through a duplication event and then lose their original host function over time whilst retaining their ability to bind the pathogen effector, examples include the psuedokinases RKS1 and ZED1 which have no kinase activity but are required to cause resistance to *Xanthomonas campestris* and *Pseudomonas syringae* respectively (Van Der Hoorn and Kamoun, 2008; Roux et al., 2014; Wang, G. et al., 2015; Kourelis, van der Hoorn and Sueldo, 2016). Domains of decoy proteins can subsequently be incorporated into the architectures of NLRs enabling to directly perceive the interaction of decoy domain with pathogen effectors (Kroj et al., 2016).

WRR5B and *WRR7* both encode the integrated LIM-peptidase decoy domain and are both present in the *A. thaliana* Col-0 genome. Additionally, resistance genes encoding LIM-Peptidase domains have been identified in other species, including species outside of the Brassicaceae family such as *Malus domestica* and *Cicer arietinum* (Srivastava and Verma, 2015; Sarris, P. F. et al., 2016). This raises the interesting prospect that LIM-Peptidase encoding resistance genes can provide resistance against multiple phytopathogens because *A. candida* is a specialist phytopathogen of Brassicaceae species (Saharan et al., 2014). Therefore, we would expect LIM-Peptidase encoding resistance genes outside of the Brassicaceae family to be potential targets for different phytopathogens.

The plant NLRome contains a diverse array of IDs that act in recognising invading pathogens and activate defence responses (Kroj et al., 2016; Sarris, P. F. et al., 2016; Van de Weyer et al., 2019). The widespread nature of integrated domains in NLRs and their prevalence in plant genomes suggests that plants have evolved a common mechanism allowing these NLR domain fusions to occur. If we wish to better inform

NLR engineering approaches that use the integrated decoy model, we need to understand the events that underpin the formation of NLRs with integrated domains. NLRs containing integrated domains are enriched in paired NLR systems suggesting that their evolution may be linked to the sensor-helper mechanism of NLR activation (Bailey et al., 2018). Due to the number of NLRs with unique integrated domains, it is likely that the NLR itself is the motile element within the genome and its insertion into the genome at various sites drives the evolution of NLR-IDs. This theory would also explain why the majority of integrated domains identified to date fall at either the C or N-termini of the NLR (Van de Weyer et al., 2019). A recent analysis, assessed the evolution of NLRs with IDs within the Poaceae family and found the CID motif (upstream of the ID) to be enriched and that the most likely mechanism of NLR-ID evolution is gene duplication and inter-chromosomal translocation by ectopic recombination (Bailey et al., 2018). Although, several transposable elements are found in NLR clusters and their role in NLR evolution has not been fully studied (Bailey et al., 2018; van Wersch and Li, 2019). It is likely that the majority of NLR-IDs evolve through the movement of NLRs themselves, it is also plausible that IDs have moved into NLRs as well.

WRR5B and WRR7 are not only an interesting model of the evolutionary mode in which NLRs gain ID fusions but are also interesting in terms of their ability to be engineered to create novel resistance using artificially selected integrated domains. To date, the majority of successful resistance gene breeding or transposition into crop lines has involved moving NLRs into the desired breeding lines without altering these resistance genes in any way (Borhan, Mohammad Hossein et al., 2010; Zhu et al., 2012; Das and Rao, 2015). However, development of these varieties is time consuming and the latent nature of resistance evolution means that single gene varieties are unlikely to be durable in the field and that to provide longer term resistant varieties, stacking of R-genes is going to be necessary to combat crop disease (Zhang, M. and Coaker, 2017). Recently, non-host resistance in *Arabidopsis* populations has been determined to be caused by a multitude of NLRs that are differentially present in various combinations in the genomes of *A. thaliana* ecotypes (Cevik et al., 2019). Therefore, in this system a species has evolved to have a genetic pool of resistance genes conferring resistance to the same pathogen resulting in mixed populations with highly fluid pools of NLRs that provide broad resistance to

the invading pathogen. For crop cultivars to be durable we will have to emulate this natural model in crop species. In order to be able to stack R-genes successfully, a large number of R-gene and R-gene derivatives are going to have to be produced in order to maintain an arsenal of R-genes that are able to be deployed against phytopathogens. Therefore, having a deep understanding of how R-genes evolve will be informative in generating novel NLRs containing ID fusions.

WRR5B and *WRR7* are resistance genes from two distinct resistance gene classes, the TNLs and the CC_Rs. Therefore, the integration of the LIM-Peptidase domain into the architectures of these NLRs must have occurred in at least two independent events. The presence of other LIM-Peptidase domains in species outside of the Brassicaceae family suggests that these events could be quite old. Therefore, in this chapter we performed analysis on all the known plant resistance genes encoding LIM-Peptidase domains in order to determine how many times this fusion event has occurred, and we attempt to narrow down the time frame of when these events happened. Moreover, we attempt to re-create the possible evolutionary event that led to the evolution of *WRR7*. To do this, we used domain swapping techniques to exchange highly similar LIM-peptidase domains with the WRR7 LIM-peptidase domain to determine whether this approach is a viable option in engineering novel resistance genes with novel integrated domains.

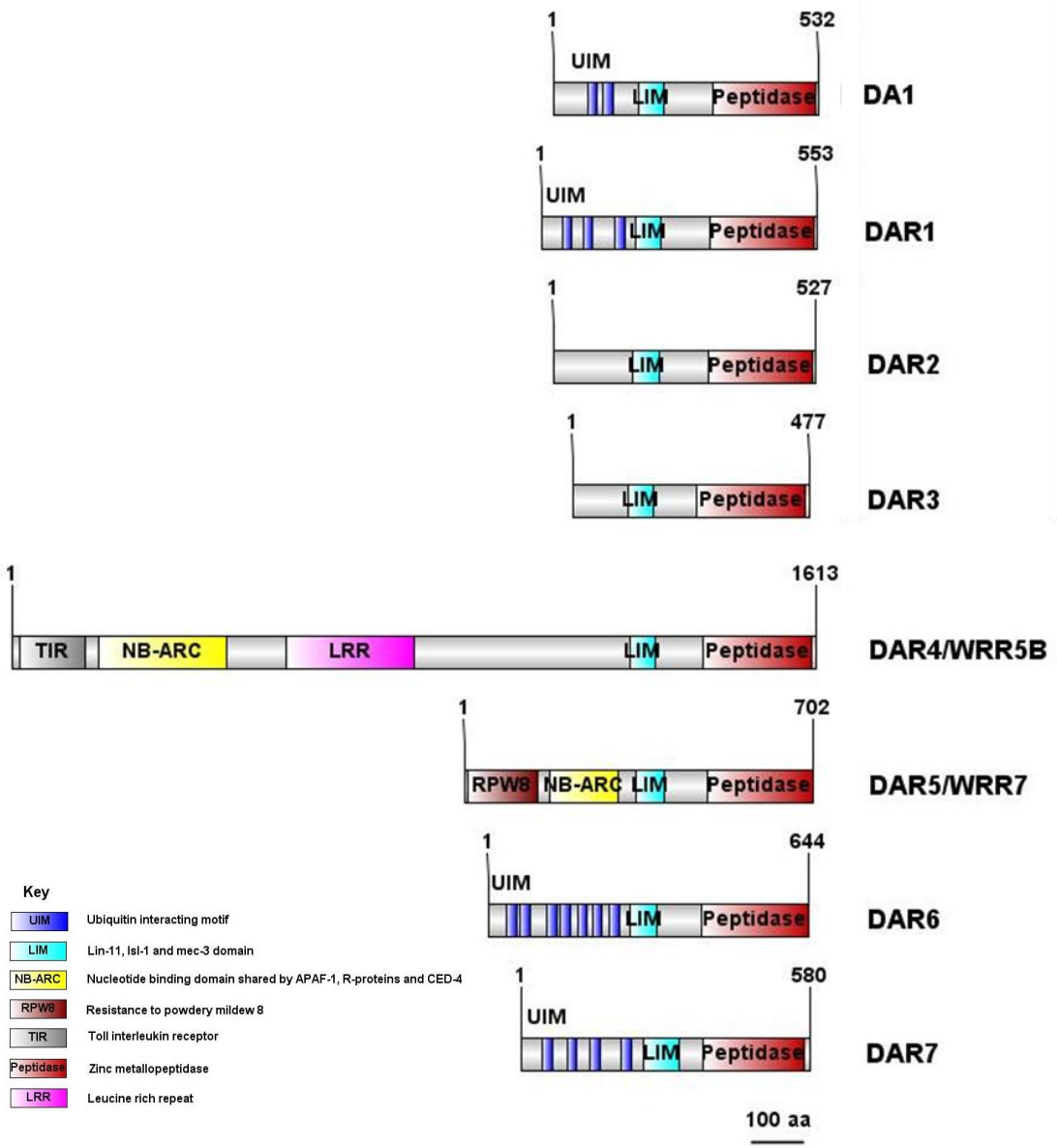


Figure 6.1: The *Arabidopsis thaliana* DA1 family proteins
Schematic representation of *Arabidopsis thaliana* DA1 family proteins with their associated domains. Proteins are numbered from the start to stop codon. Domain locations were predicted using SMART (Letunic, Ivica, Doerks and Bork, 2015). Scale bar represents 100 amino acids.

Results

The DA1 family is divided into two distinct clades

Both *WRR5B* and *WRR7* encode an integrated LIM-Peptidase domain and confer resistance against *A. candida* isolate AcEM2 in *A. thaliana*. The presence of these domains mean they fall into the DA1 protein family which is characterised by the presence of the LIM and Zinc metallopeptidase domains in this sequential arrangement (Fig 6.1). In *A. thaliana* there are eight members of the DA1 family, DA1 as well as DAR1-7 including WRR5B (DAR4) and WRR7 (DAR5). Most of the DA1 family proteins are comprised of the LIM-Peptidase domain as well as a number of ubiquitin binding motifs apart from the two resistance genes which also encode domains associated with NLRs. The presence of two resistance genes that encode integrated domains from one protein family, that are able to provide resistance independently of one another against the same pathogen, suggests that this domain is a putative target of pathogen effectors. Therefore, we hypothesised that an AcEM2 effector targets the LIM-peptidase domain of one of the DA1 family members and this interaction also occurs with the LIM-peptidase domain of WRR5B and WRR7. We also hypothesised that the interaction of the *A. candida* effector with the integrated LIM-Peptidase domain is highly specific and the target of the *A. candida* effector would share close homology with the integrated domain of WRR5B and WRR7. To dissect this hypothesis, we performed sequence analysis of the DA1 family LIM-Peptidase domains in *A. thaliana* to identify which of the domains of the DA1 family proteins was closest in homology to the LIM-peptidase domains of WRR5B and WRR7.

We found that there are two distinct clades of DA1 family proteins based on sequence analysis of the LIM-peptidase domains (Fig 6.2). Clade I contains three members DA1, DAR1 and DAR2, whilst Clade II contains the remainder of the DA1 family proteins (DAR3-7), including both WRR5B and WRR7 (Fig 6.2). We found that the WRR7 integrated LIM-peptidase domain is closest in homology with the DAR6 LIM-peptidase domain and that the integrated domain in WRR5B shows closest homology to the LIM-Peptidase domains of DAR3 and DAR7, although the support for either scenario was low (Fig 6.2). Our results confirm previous findings by (Peng, Yuancheng et al., 2015) that there are two distinct clades within this protein family.

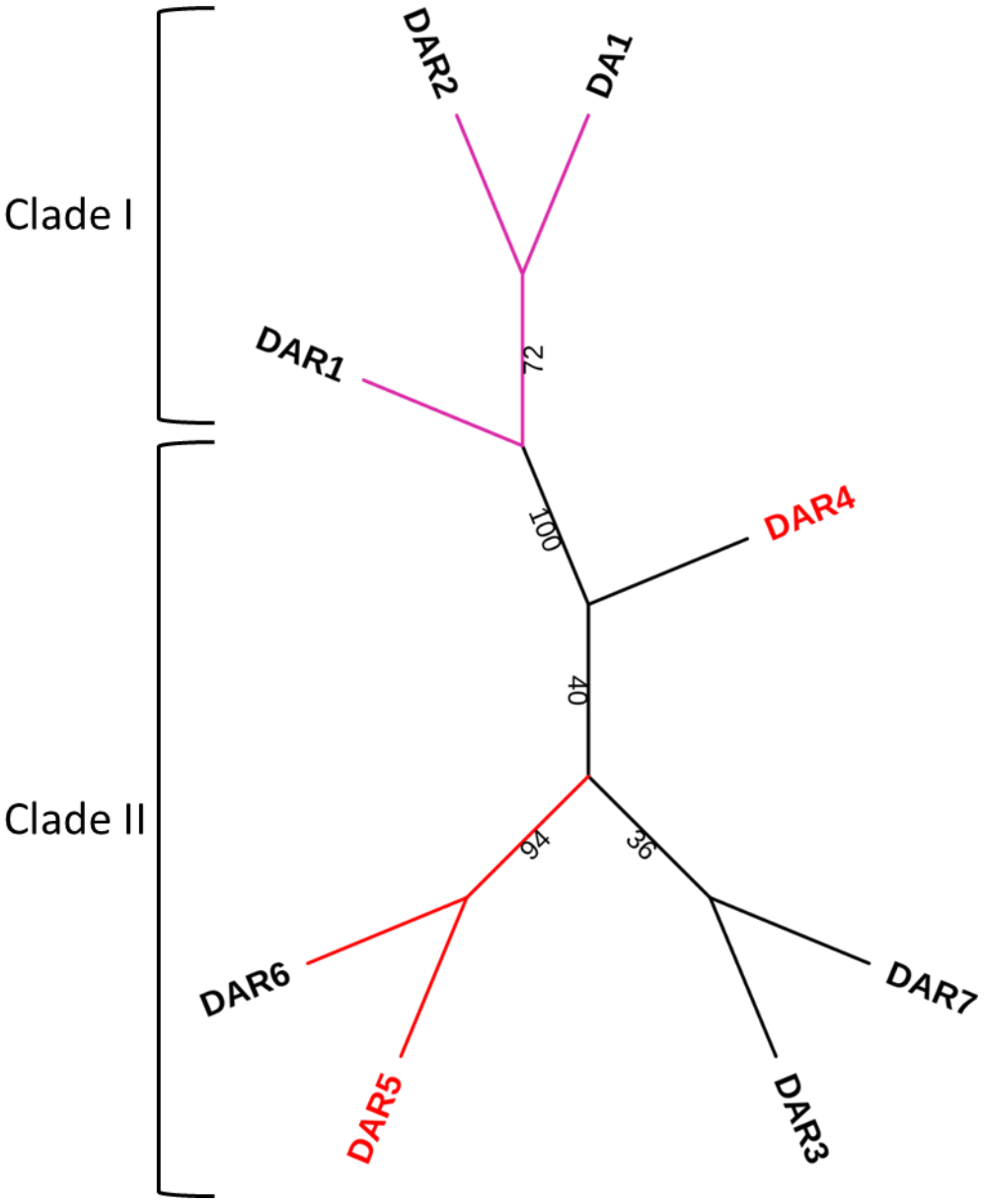


Figure 6.2: WRR7 and WRR5B are part of Clade II of the DA1 family
Maximum likelihood (Phy ML) star tree of the *Arabidopsis thaliana* DA1 family proteins. The star tree was built using amino acid sequences from each proteins from the start of the LIM domain to the C-terminal end of each protein and aligned using MUSCLE and Phy ML analysis was performed in SeaView (Edgar, 2004). Bootstrap values were calculated using 100 replicates. WRR5B (DAR4) and WRR7 (DAR5) leaf labels are in Red, non-resistance gene leaf labels are in black. Clade I DA1 proteins have pink branches whereas Clade II proteins have black/red branches.

Overexpression of DA1 family proteins in *Arabidopsis thaliana* results in phenotypic abnormalities

Our sequence analysis suggested that the LIM-Peptidase domains of WRR5B and WRR7 are closest in homology with the Clade II DA1 family members (Fig 6.2). Therefore, it is most likely that *A. candida* effectors are targeting Clade II DA1 family members during infection, particularly DAR6 which shares close homology with the LIM-Peptidase domain from WRR7. To determine whether *A. candida* infection affects DA1 family proteins, we first needed to generate *A. thaliana* lines overexpressing DA1 family members. Therefore, we generated T_3 overexpression lines of DA1 family members (excluding *WRR5B* and *WRR7*) in the AcEM2 susceptible *A. thaliana* ecotype Ws-2. Each DA1 family protein was overexpressed using the 35S Cauliflower mosaic virus promoter and C-terminally epitope tagged with the His-FLAG tag.

It has already been documented that mutating DA1 family members causes phenotypic changes to *A. thaliana* plants for example *da1*, *dar1* and *dar2* mutants show inhibited growth phenotypes in both *A. thaliana* and *Brassica napus* (Li, Yunhai et al., 2008; Peng, Yuancheng et al., 2015; Wang, J.-L. et al., 2017) and overexpression of DA1 also inhibits growth compared to wild type plants (Vanhaeren et al., 2017). However, no previous phenotypic characterisation has been performed on lines with altered DAR3, DAR6 or DAR7 protein levels. Therefore, we phenotypically analysed all the non-resistance gene DA1 family proteins in *A. thaliana* Ws-2 plants to determine any phenotypic defects that could be caused by the overexpression of these proteins.

We found that overexpression of *DA1-HF* and *DAR1-HF* caused a reduction in growth after four weeks and that overexpression of *DAR1-HF* induced early flowering in short day conditions after just 4 weeks of growth compared to control lines overexpressing *GUS-HF* (Fig 6.3). There was little difference between lines overexpressing *DAR3-HF* and *DAR7-HF* compared to transgenic control lines overexpressing *GUS-HF* (Fig 6.3). However, overexpression of *DAR6-HF* caused substantial phenotypic defects including a reduction in growth, narrow leaves, malformation of the rosette, delayed flowering and no primary bolt production (Fig 6.3 and 6.4). Therefore, DAR6 plays an important role in controlling developmental programmes in *A. thaliana*.

Figure 6.3: Ws-2 lines overexpressing *DA1* family genes display phenotypic abnormalities

Phenotypes of independent homozygous T_3 lines overexpressing non-resistance *DA1* family genes using the 35S cauliflower mosaic virus promoter (35S) and epitope tagged with His-FLAG (HF), T_3 lines in *Arabidopsis thaliana* Ws-2 background. Images taken four weeks post sowing and the scale bar represents 1cm.

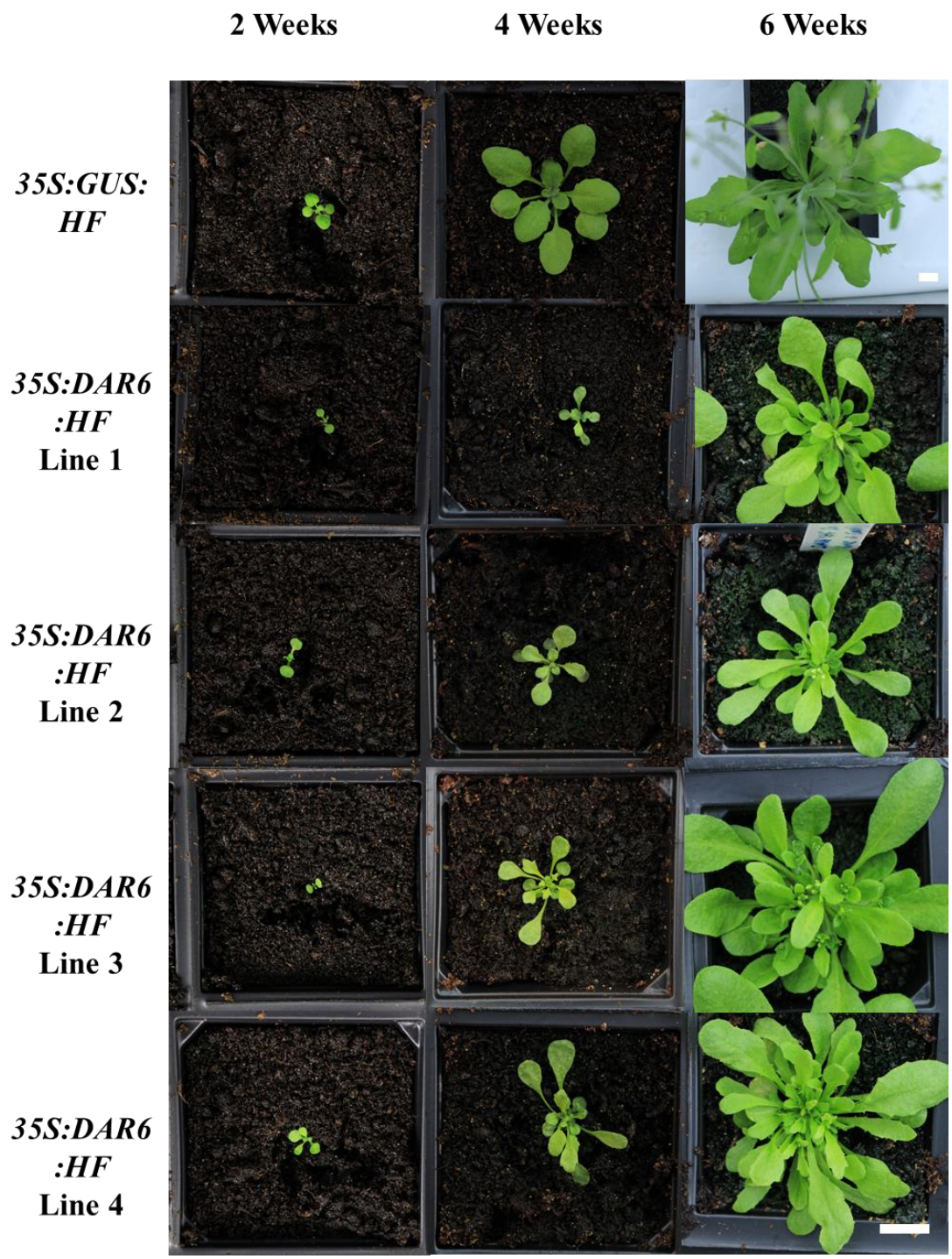


Figure 6.4: Characterisation of Ws-2 lines overexpressing *DAR6:HF*
Phenotypes of independent homozygous *Arabidopsis thaliana* T_3 lines overexpressing *DAR6:HF* and *GUS:HF* driven by the 35S cauliflower mosaic virus promoter (35S) and epitope tagged with His-FLAG (HF). Images taken 2, 4, and 6 weeks after sowing. Scale bars represent 1cm.

DA1 family proteins are unstable following *A. candida* infection

The LIM-peptidase domain that categorises the DA1 family proteins is present in the architecture of both WRR5B and WRR7. Both of these LIM-peptidase domain encoding resistance genes are able to confer resistance to *A. candida* isolate AcEM2. Integrated decoy domains have previously been shown to interact with pathogen effectors, such as the WRKY domain of RRS1 that binds to *Ralstonia solanacearum* effector PopP2 and *Pseudomonas syringae* effector AvrRPS4 (Sarris, Panagiotis F. et al., 2015). Therefore, the presence of the LIM-Peptidase domain in two resistance genes conferring resistance to *A. candida* indicates that this pathogen contains effector(s) that potentially target the LIM-Peptidase domain containing proteins during infection. To determine whether *A. candida* affected the protein stability of DA1 family proteins during infection, we cloned and overexpressed all the DA1 family members (excluding *WRR5B* and *WRR7*) and epitope tagged each protein with His-FLAG tag. We then inoculated each line with *A. candida* isolate AcEM2 or mock inoculated them with water, 4 weeks post germination and performed immunoblot analysis on independent lines to determine whether we could detect any changes in the protein levels.

We found that we could detect depletions in the protein levels of all DA1 family members in *A. thaliana* Ws-2 plants after infection with AcEM2, apart from DAR2 lines where protein levels were too low to detect and that there was no observable depletion of the GUS-HF control after AcEM2 infection (Fig 6.5 and Table 6.1). However, this depletion in protein levels was not consistent and therefore hard to draw conclusions from. The most depleted proteins that we identified following AcEM2 infection were DAR3 and DAR6 although we had limited lines to test for this depletion (Table 6.1).

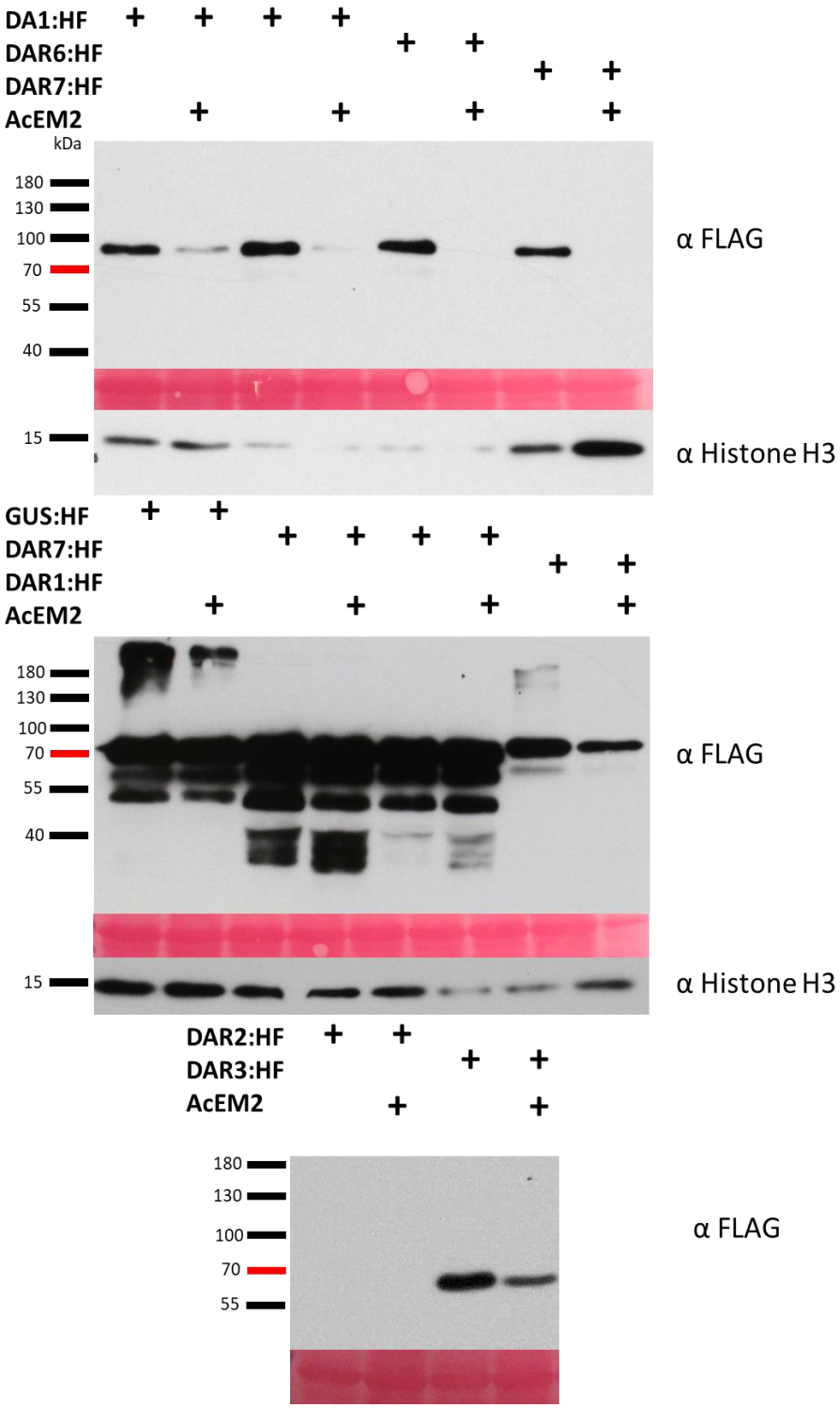


Figure 6.5: Immunoblots of DA1 family proteins following *Albudo candida* infection

Immunoblot results of total protein extracts from the leaves of 4-week-old *Arabidopsis thaliana* Ws-2 plants overexpressing His-Flag epitope tagged DA1 family proteins. Proteins were extracted 4 days post inoculation with *Albugo candida* isolate AcEM2 or mock inoculation with water. Proteins were visualised using α-FLAG antibody. Pink tiles show Ponceau-S stained membranes and bottom immunoblot shows samples visualised using α-histone H3 antibody.

Table 6.1: Summary of DA1 family immunoblot results after infection with *Albugo candida*

Summary of immunoblot results from *Arabidopsis thaliana* Ws-2 plants overexpressing His-FLAG epitope tagged *DA1* family members 4 days post inoculation with *Albugo candida* isolate AcEM2 or mock inoculated with water. Replicates were recorded where protein levels were noticeably depleted between mock inoculated and AcEM2 inoculated plants after AcEM2 infection.

DA1 family member overexpressed	Change in protein level detected	Number of replicates	Number of replicates depletion detected in	Percentage of replicates showing depletion
DA1	Yes	14	3	21%
DAR1	Yes	6	3	50%
DAR2	N/A	2	N/A	N/A
DAR3	Yes	3	3	100%
DAR6	Yes	4	3	75%
DAR7	Yes	10	3	30%

The DA1 family Clade II proteins as decoys

There are currently no known functions of the clade II DA1 family proteins apart from the involvement of WRR5B and WRR7 in conferring resistance to *A. candida* (Fig 3.5 and 3.10) and our finding that DAR6 is involved in regulating plant development in *A. thaliana* when constitutively overexpressed (Fig 6.3 and 6.4). We did not find any phenotypic effects of overexpressing DAR3 or DAR7 in *A. thaliana* Ws-2 plants (Fig 6.3). Therefore, we hypothesised that DAR3 and DAR7 could be acting as decoy proteins that could recognise an *A. candida* effector. Both of these proteins contain CAMTA binding motifs, in DAR7 the motif is 290 bp upstream of the ATG and in DAR3 the motif lies 5 bp downstream of the ATG. In addition, DAR7 showed differential expression following *A. candida* infection and was part of the cluster of genes whose expression profiles were enriched for CAMTA transcription factor binding motifs (Fig 5.4 and Table 5.1). Therefore, both of these genes could be co-regulated with *WRR7* by CAMTA2 following *A. candida* infection. Any protein acting as a decoy would be under reduced selective pressure as decoy proteins divest themselves of their original host functions (Roux et al., 2014). Therefore, we analysed the active site motif (HEMMH) of the peptidase domains of DA1 family proteins to see whether this domain has been affected indicating a loss of function of this domain. As well as motif analysis, we theorised that a signature of the evolution of a decoy protein would be an increase in the mutation rate in these proteins compared to the functional host protein that they are mimicking. To analyse whether clade II DA1 family proteins are evolving at a faster rate than their clade I counterparts, we performed maximum likelihood phylogenetic analysis using the LG model of evolution to predict the rate of genetic change (Le and Gascuel, 2008).

We found that DAR3 was the only DA1 family member that had no intact HEMMH peptidase motif (Fig 6.6). We also found that DAR7 has a 27 amino acid insertion immediately upstream of the active site motif which could impair its function (Fig 6.6). In addition, our evolutionary analysis revealed that *DAR3, DAR6* and *DAR7* along with their Brassicaceae homologs were all evolving at a faster rate than *DA1, DAR1* and *DAR2* and their respective homologs in the Brassicaceae (Fig 6.7), suggesting that *DAR3, DAR6* and *DAR7* are all under reduced selective pressure compared to *DA1, DAR1* and *DAR2* and could be functioning as decoy proteins.

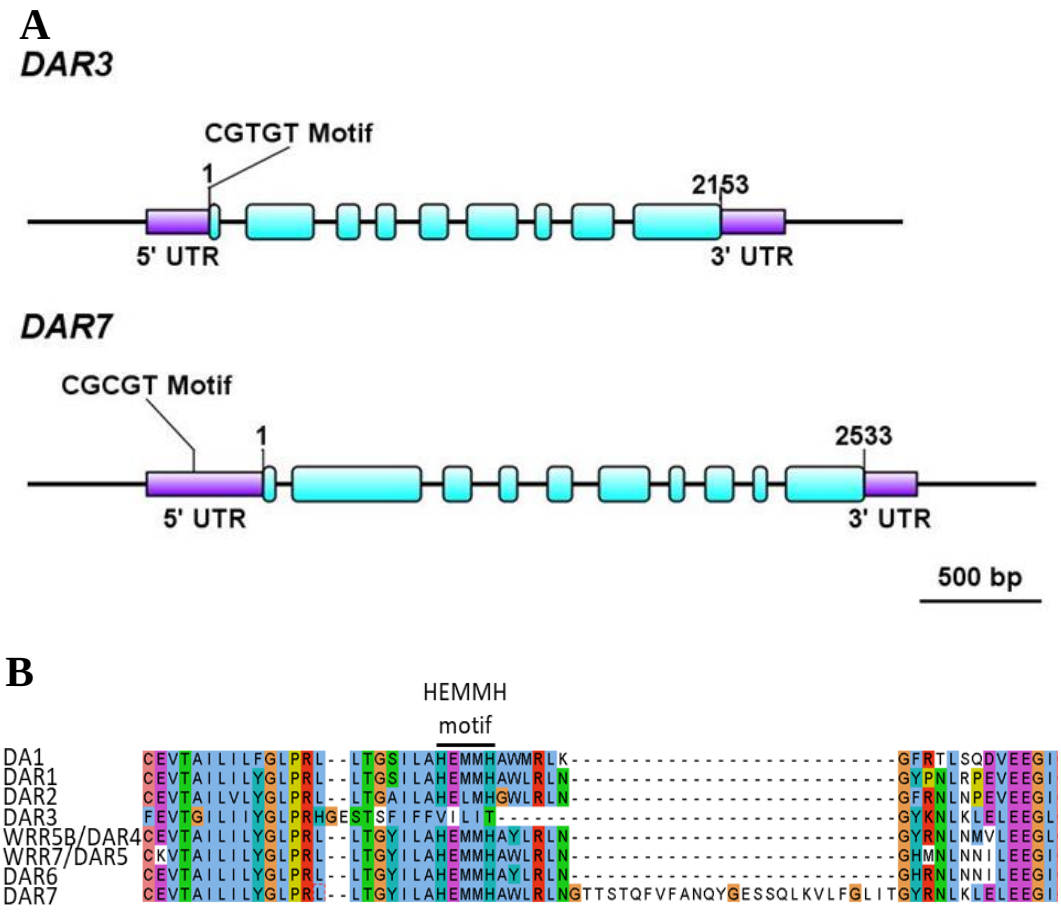


Figure 6.6: Alignment of Peptidase active site motif
A- Schematic diagram of the nucleotide sequences of *DAR3* and *DAR7*, showing the position of their CAMTA binding motifs. The genes are numbered from the start codon to the stop codon, exons are represented in blue.
B- Muscle alignment of DA1 family peptidase domain amino acid sequence showing the HEMMH active site motif (Edgar, 2004). Amino acids are coloured using the Clustal colour scheme and visualised in Jalview software (Waterhouse et al., 2009).

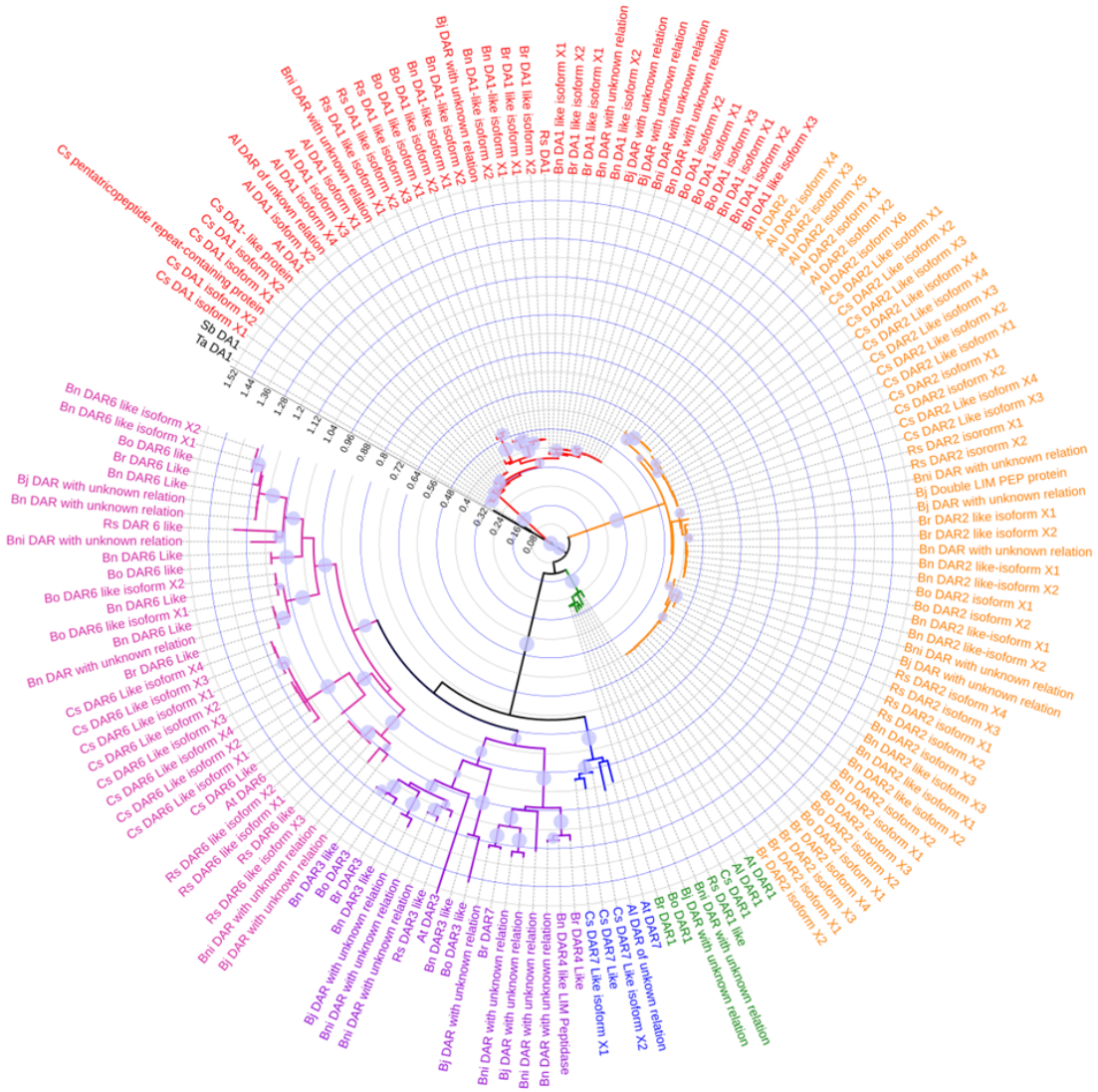


Figure 6.7: Brassicaceae Clade II DA1 family members are evolving faster than Clade I family members

Phy Maximum likelihood analysis using the LG model of evolution of non-NLR DA1 family members and their respective homologues identified from Brassicaceae species that have had their full genomes sequenced: *A. thaliana* (At), *A. lyrata* (Al), *C. sativa* (Cs), *B. juncea* (Bj), *B. napus* (Bn), *B. nigra* (Bni), *B. oleracea* (Bo) and *R. sativus* (Rs). Phylograms were constructed using amino acid sequences from the start of the LIM domain to the end of the protein sequence aligned using MUSCLE (Edgar, 2004). Monocot DA1 sequences from *Sorghum bicolor* (Sb) and *Triticum aestivum* (Ta) were used as the outgroup. Branch lengths represent the amount of genetic change between the aligned sequences and circles correspond to bootstrap values >80 following 100 replicates. Branches coloured red depict annotated DA1 sequences and their homologues, dark green for DAR1, yellow for DAR2, purple for DAR3, pink for DAR6 and dark blue for DAR7. The tree was annotated in iTOL (Letunic, I. and Bork, 2016).

LIM-Peptidase domains of WRR5B and WRR7 do not show peptidase activity on EOD1

Both WRR5B and WRR7 contain integrated LIM-Peptidase domains in their architecture. The presence of these integrated domains in both proteins suggests that there are similarities in how these two resistance genes function, even though WRR5B belongs to the TNL class of resistance genes and WRR7 belongs to the CC_R family of resistance genes. Integrated decoy domains are believed to lose their original function in the process of becoming incorporated into NLRs. However, we identified that the active site of the peptidase domain in both WRR5B and WRR7 is intact (Fig 6.6). Therefore, this would suggest that these domains have retained their original host function as active peptidases.

The LIM-Peptidase domains are found in the DA1 family of proteins in plants which are involved in cell growth and development. The peptidase domains of DA1 and DA1 family members such as DAR1 cleave E3 ligase Enhancer of DA1 (EOD1) (Dong, H. et al., 2017). The active site of the peptidase domain is known to be the conserved HEMMH motif, both the LIM-Peptidase domains from WRR5B and WRR7 contain an intact active site motif and could therefore cleave EOD1 (Fig 6.6). To test whether the LIM-Peptidase domains have retained or lost their peptidase activity in the process of integration into an NLR, we fused the N-terminal region of DA1 (up to its LIM domain) with the LIM-Peptidase domain of DAR1, DAR4 (WRR5B) and DAR5 (WRR7) (Fig 6.8). The N-terminal region of DA1 was used as this protein has previously been shown to cleave EOD1 (Dong, H. et al., 2017). Therefore, fusing the LIM-Peptidase domains from WRR5B and WRR7 to the DA1 N-terminus should show whether these domains have retained their ancestral function, the LIM-Peptidase domain from DAR1 was also fused to the DA1 N-terminus as a positive control. These constructs were then overexpressed by the use of the 35S cauliflower mosaic virus promoter and C-terminally tagged with a His-Flag tag and transiently co-expressed with *A. tumefaciens* GV3101 strains overexpressing V5 epitope tagged EOD1 in *N. benthamiana*.

As shown previously, DA1-HF and DAR1-HF constructs effectively cleaved EOD1 (Dong, H. et al., 2017), and the fusion proteins made up of the DA1 N-terminal end fused to the DAR1 LIM-Peptidase domain was also able to cleave EOD1 (Fig 6.9). However, we found that the LIM-Peptidase domains from WRR5B and WRR7 fused to the N-terminal end of DA1 were unable to cleave EOD1 (Fig 6.9). Our findings support the hypothesis that integrated domains lose their ancestral function in the process of becoming incorporated into a resistance gene and that the LIM-Peptidase domains of WRR5B and WRR7 are both acting as decoy domains.

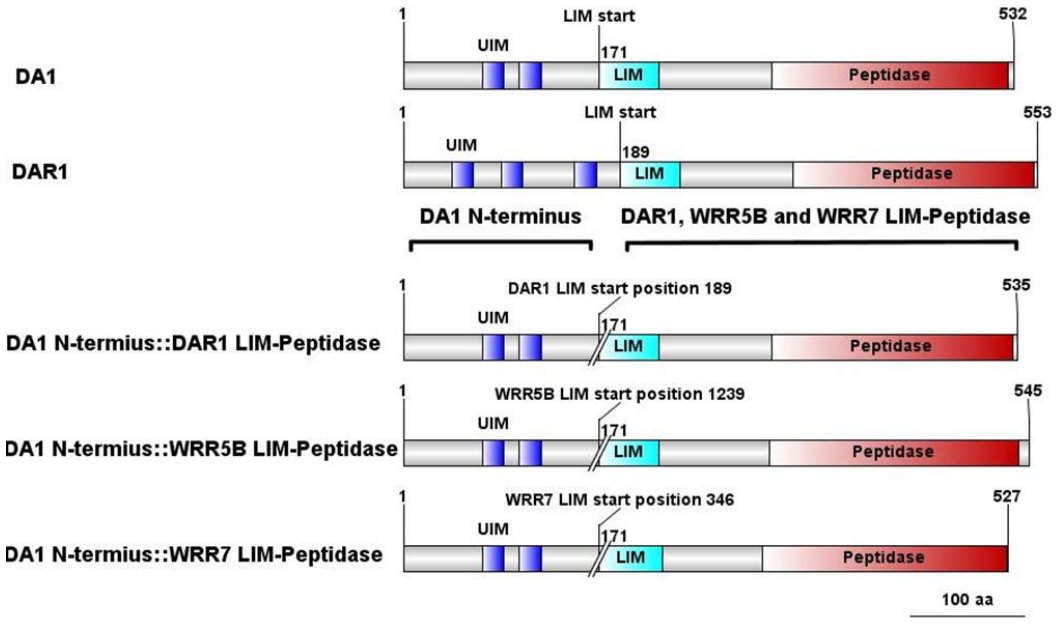


Figure 6.8: DA1 LIM-Peptidase domain swaps with DAR1, WRR5B and WRR7

Schematic representation of the DA1 and DA1 family LIM-Peptidase domain swaps. The DA1 LIM-peptidase domain was swapped for the LIM-peptidase domains from DAR1, WRR5B/DAR4, and WRR7/DAR5 proteins. Domain predictions were performed using SMART (Letunic, Ivica, Doerks and Bork, 2015). Scale bar shows 100 amino acids.

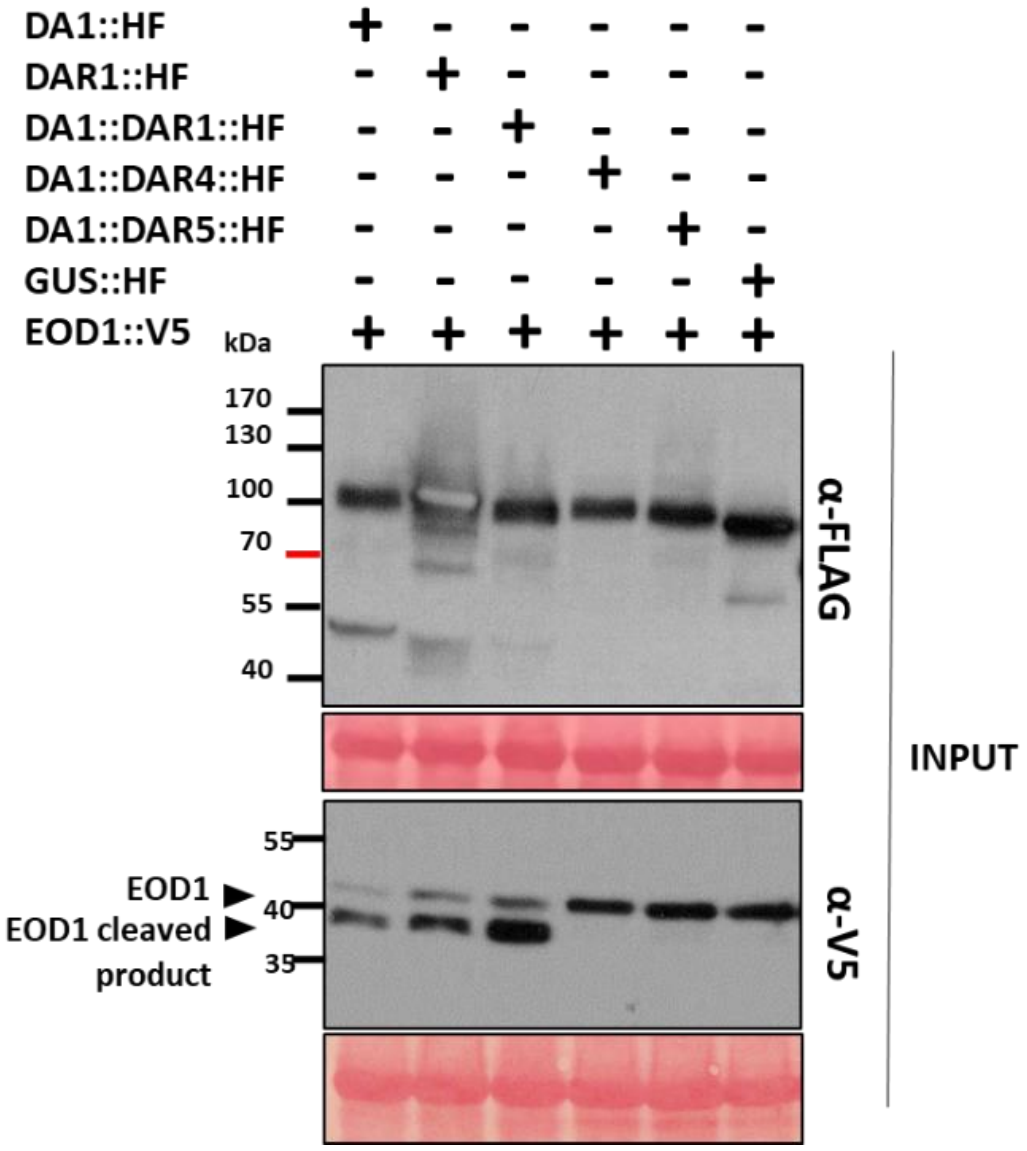


Figure 6.9: LIM-Peptidase domains of WRR5B and WRR7 are unable to cleave EOD1

Immunoblot of overexpressed, His-FLAG epitope tagged DA1, DAR1 and DA1 N-terminal fusion constructs with the LIM-Peptidase domains of DAR1, WRR5B (DAR4) and WRR7 (DAR5) proteins. Cloned constructs were transformed into *Agrobacterium tumefaciens* strain GV3010 and transiently in *Nicotiana benthamiana* leaves. Proteins were extracted 3 days post infiltration. Pink tiles show the Rubisco protein after Ponceau-S staining of the membranes.

Identification of Resistance genes with integrated LIM-Peptidase domains

It has been previously documented that there are additional NLRs in the plant kingdom that contain integrated LIM-Peptidase domains, including in species that are not part of the Brassicaceae family (Srivastava and Verma, 2015; Sarris, P. F. et al., 2016). We have shown that the LIM-Peptidase encoding resistance genes *WRR5B* and *WRR7* confer resistance to Brassicaceae specialist pathogen *A. candida.* However, the presence of LIM-Peptidase domain encoding resistance genes in other plant families suggests that these resistance genes could confer resistance to other phytopathogens. To explore the diversity of resistance genes with integrated LIM or Peptidase domains, we compiled a list of these known LIM-Peptidase domain encoding resistance genes and ran BLASTP searches in NCBI, LIS, BRAD, GDR and the Sol genomics network to identify any other resistance genes that encode integrated LIM-Peptidase domains.

We were able to identify LIM-peptidase containing NLRs from fifteen species (Fig 6.10 and 6.11), 8 in the Brassicaceae family including *A. thaliana* (Fig 6.10) as well as six that were present in species outside of the Brassicaceae family (Fig 6.11) including species from the Rosaceae, Vitaceae and Fabaceae families (*Malus domestica, Cicer arietinum, Prunus persica, Prunus meme, Vitis vinifera and Medicago truncatula*). There were no LIM-peptidase domain encoding resistance genes identified in the Solanaceae (Fig 6.11). Some of the species identified that harbour resistance genes encoding LIM-peptidase domains such as *P. persica* and *C. arietinum* harboured multiple such resistance genes in their genomes (Fig 6.11). The presence of a common integrated domain in four plant families suggests that this type of integrated domain fusion is able to confer resistance to multiple plant pathogens and that the LIM-peptidase domain may be a common target for virulence genes of multiple phytopathogens.

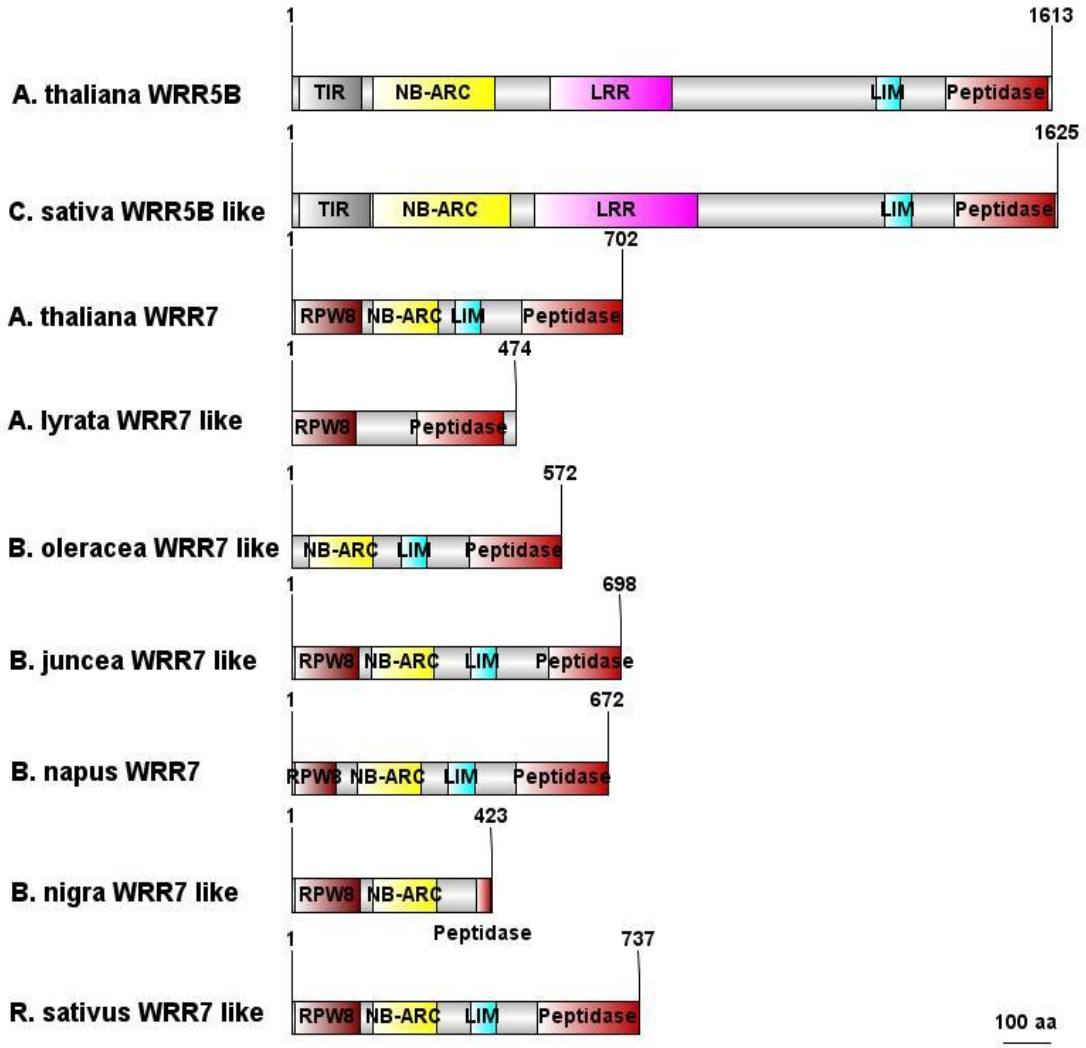


Figure 6.10: Brassicaceae resistance proteins containing integrated LIM or Peptidase domains

Protein schematics of resistance genes in the Brassicaceae family containing LIM-Peptidase domains. Proteins were identified using BLASTp searches of all the available Brassicaceae genomes in NCBI and the Brassicaceae genome resource (BRAD). Resistance genes encoding LIM-Peptidase domains were identified in *Arabidopsis thaliana*, *Camelina sativa*, *Arabidopsis lyrata*, *Brassica oleracea*, *Brassica juncea*, *Brassica napus*, *Brassica nigra and Raphanus sativus*. Domain predictions were performed using SMART predictions (Letunic, Ivica, Doerks and Bork, 2015) and the scale bar represents 100 amino acids.

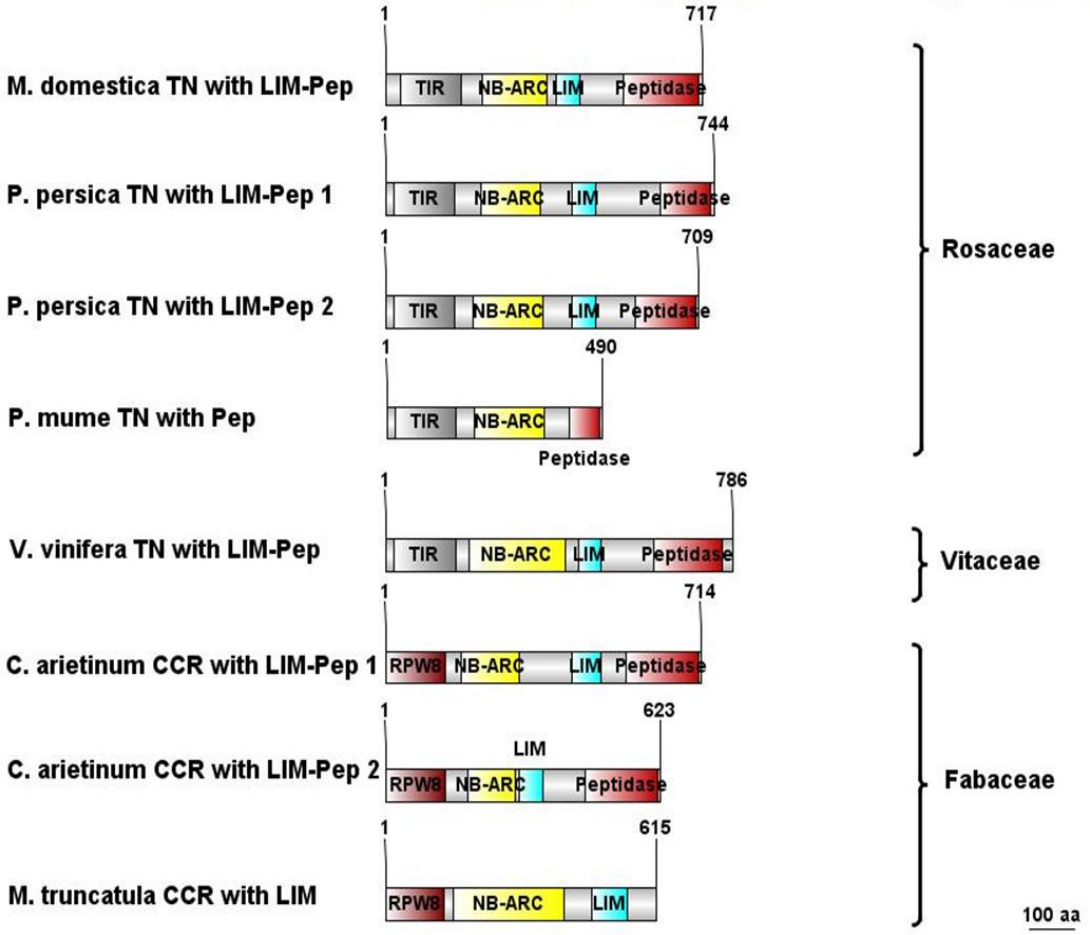


Figure 6.11: Non-Brassicaceae Resistance proteins contain integrated LIM-peptidase domains

Protein schematics of resistance proteins outside of the Brassicaceae family containing LIM-Peptidase domains. Proteins were identified using BLASTp searches of all the available plant genomes in NCBI as well as the legume information service, the genome database for Rosaceae and the Sol genomics network. Resistance genes encoding LIM-Peptidase domains were identified in *Malus domestica, Prunus persica, Prunus mume, Vitis vinifera, Cicer arietinum and Medicago truncatula* genomes. Domain predictions were performed using SMART predictions (Letunic, Ivica, Doerks and Bork, 2015) and the scale bar represents 100 amino acids.

Clade II of the DA1 family evolved at the base of the Brassicaceae lineage

To determine the evolutionary origin of the integration of the LIM-peptidase into WRR7 and WRR5B, we mined the genomes of Brassicaceae species getting progressively more diverged from *A. thaliana* for any LIM-Peptidase containing proteins. The amino acid sequence of the LIM-Peptidase domain from identified proteins was then used to create a maximum likelihood (RaxML) phylogenetic tree. We also incorporated some of the peptidase sequences from resistance genes encoding LIM-peptidase integrated domains from non-Brassicaceae species (Fig 6.11) into the tree, to determine how many independent evolutionary events could have resulted in NLRs containing an integrated LIM-peptidase domain.

We found that clade II DA1 family proteins (DAR3-DAR7) were only present in species in the Brassicaceae family and were not present in the genomes of *Theobroma cacao* or *Carica papaya*, the two genomes most closely related to the Brassicaceae family that we analysed (Fig 6.12). Therefore, the evolution of this group of proteins likely occurred in one single duplication event during the early diversification of the Brassicaceae family. All of the clade II members of this family with the exception of WRR5B (DAR4) are located at a single locus on chromosome 5 suggesting that these proteins share a common origin. WRR5B is most closely related to the Clade II DA1 proteins (Fig 6.1). Therefore, it is likely that the WRR5B LIM-Peptidase domain originated from this clade and later translocated into the WRR5 locus.

The LIM-Peptidase encoding resistance genes found inside and outside the Brassicaceae contained resistance genes in both the TNL and CC_R classes of resistance genes (Fig 6.10 and 6.11). The *A. thaliana* LIM-peptidase containing NLRs are both part of Clade II of DA1 family proteins that evolved in the Brassicaceae lineage (Fig 6.12). However, we identified LIM-peptidase containing NLRs from the Rosaceae and Fabaceae families as well (Fig 6.11). The LIM-Peptidase domain containing NLRs identified from *Prunus persica*, *Cicer arietinum* and *Vitis vinifera* all shared closer homology with DA1 family clade I proteins or were part of a separate DA1 family clade that is not present in the Brassicaceae genomes tested (Fig 6.12). Therefore, it is likely that an integration event between NLRs and LIM-Peptidase domains happened on at least four separate occasions,

twice during the evolution of the Brassicaceae family and at least two times outside of the Brassicaceae family.

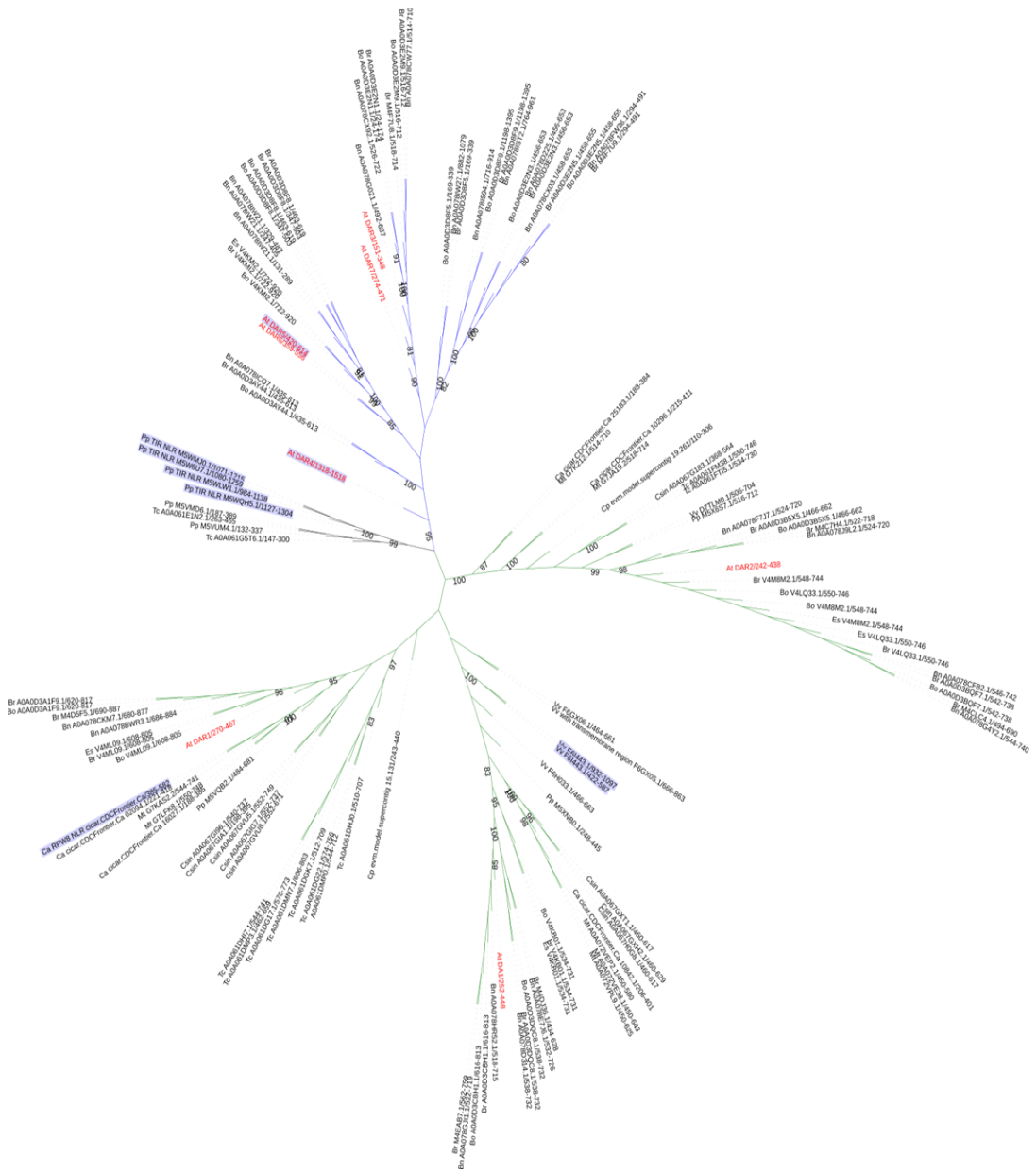


Figure 6.12: DA1 family Clade II evolved at the base of the Brassicaceae lineage

Maximum likelihood (RaxML) analysis of peptidase (pfam: PF12315) sequences from selected Brassicaceae species as well as closely related species from just outside the Brassicaceae lineage (*Theobroma cacao* and *Carica papaya*) and *Prunus persica, Cicer arietinum and Vitis vinifera* resistance genes encoding LIM-peptidase domains. RaxML was performed using the LG model of evolution and bootstraps with values >80 of 100 replicates are shown. Clade I proteins branches are shown with green branches and Clade II with blue branches, resistance gene leaves are highlighted in blue and *Arabidopsis thaliana* DA1 family leaf labels are in red.

The *Camelina sativa* genome harbours a homolog of *WRR5B* but not *WRR7*

Earlier we showed that the clade II DA1 family members evolved after the divergence of the Brassicaceae family (Fig 6.12). However, we wanted to track the origin of *WRR5B* and *WRR7* more finely. To accomplish this, we ran BLASTP searches of *WRR5B* and *WRR7* amino acid sequences against genomes within the Brassicaceae family using NCBI and the Brassica database (BRAD) (Cheng et al., 2011).

We identified seven additional LIM-peptidase encoding resistance genes in Brassicaceae genomes other than *A. thaliana WRR5B* and *WRR7* (Fig 6.10). *WRR7* homologs are present throughout the Brassicaceae family, whereas we could only identify a *WRR5B* homolog in *Camelina sativa* (Fig 6.10). Interestingly, the *C. sativa* genome harboured an extra Non-NLR LIM-Peptidase domain containing protein which was closest in similarity with the LIM-Peptidase domain of *WRR5B* (69% identity match) as well as a homolog of *WRR5B* itself which is adjacent to a homolog of *WRR5A* (Fig 6.13). We show through micro-synteny analysis using CoGE software (Haug-Baltzell et al., 2017) that the non-NLR LIM-peptidase domain located on *C. sativa* chromosome 2 is in a syntenic loci with the *A. thaliana* loci that harbours all the *A. thaliana* clade II DA1 family members except for *WRR5B,* but is not homologous with any of them (Fig 6.13). The C-terminal end of this gene shows synteny with *WRR7,* however the RPW8 and NB-ARC domains are not present in this gene (Fig 6.13). We also show that the *C. sativa WRR5B* like protein on chromosome 8 is homologous with *A. thaliana* WRR5B and both loci show synteny (Fig 6.13).

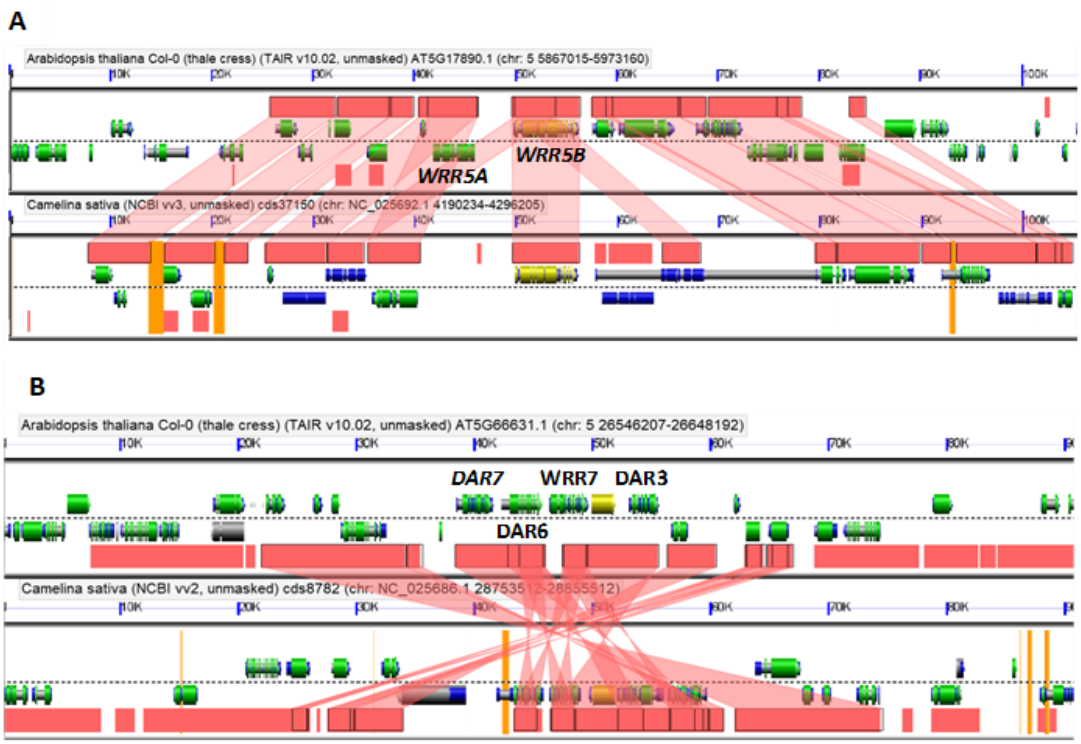


Figure 6.13: The *Camelina sativa* genome has a copy of *WRR5B* but not *WRR7*
A - Micro-synteny analysis between the *A. thaliana* Col-0 genome (top) and *Camelina sativa* (bottom) genomes at the *WRR5* loci using CoGe SynMap2 analysis (Haug-Baltzell et al., 2017). *WRR5B* is highlighted in yellow. Red blocks connected by red wedges represent syntenic regions. Genes are represented in green displaying intron and exon structure. B - Micro-synteny analysis between the *A. thaliana* Col-0 genome (top) and *Camelina sativa* (bottom) genomes at the *WRR7* loci using CoGe SynMap2 analysis (Haug-Baltzell et al., 2017). The pentricopeptide repeat containing protein just upstream of WRR7 is highlighted in yellow.

Tracking the evolution of WRR5B

We only identified homologs of *WRR5B* that also encode an integrated LIM-Peptidase domain in *Camelina sativa*. The *Camelina* genus is most closely related to the *Arabidopsis* genus therefore, it is likely that the event that led to the integration of the LIM-Peptidase domain encoding DNA into the ancestral *WRR5B* gene happened recently, prior to the divergence of these two genera. Therefore, to identify whether this was the case we analysed the *Eutrema salsugineum* genome, which is the most closely related genome two the *Arabidopsis* and *Camelina* clade (Fig 6.14), to see if we could identity a *WRR5B* homolog that does not encode an integrated LIM-Peptidase domain.

We were able to identify a *WRR5B* like gene in *Eutrema salsugineum* that sits adjacent to another NLR that bears close homology with *WRR5A* by BLASTP searching. Domain predictions of the protein sequence of the WRR5B like gene in *E. salsugineum* revealed that it does not contain an integrated LIM-Peptidase domain (Fig 6.15) and may represent a gene that bears close similarity to the ancestral *WRR5B* gene before it gained the LIM-Peptidase domain encoding DNA.

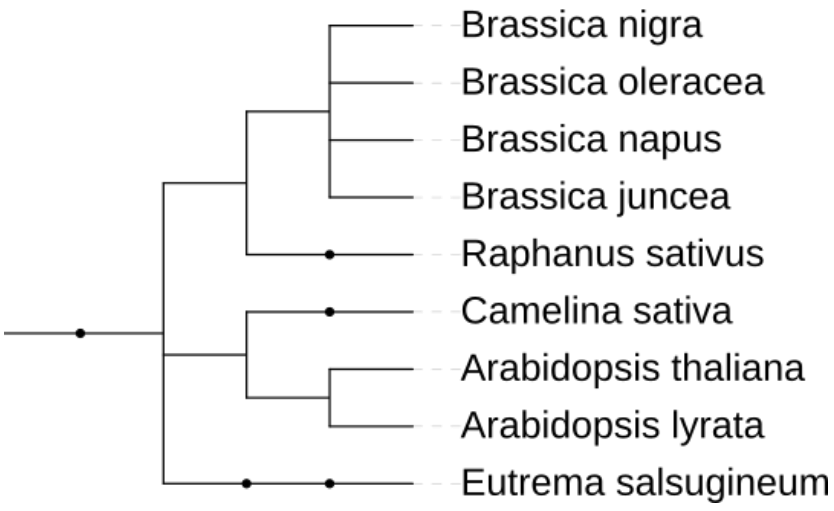


Figure 6.14: Phylogeny of Brassicaceae species with resistance genes encoding LIM-Peptidase domains
Phylogeny of the Brassicaceae species identified containing NLRs with integrated LIM-Peptidase domains as well as *E. salsugineum.* Phylogeny was drawn using PhyloT software and visualised using iTOL (Letunic, I. and Bork, 2016).

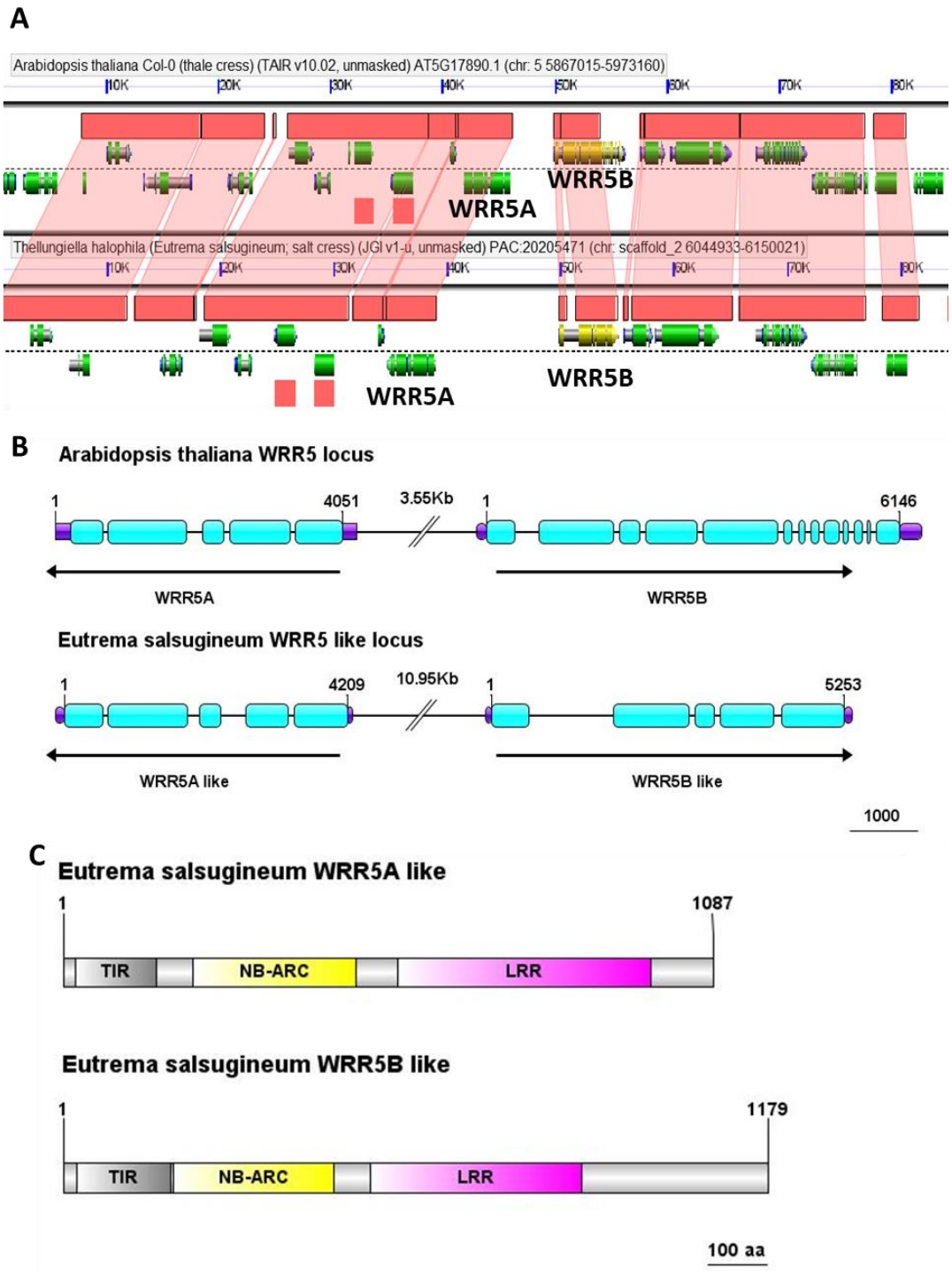


Figure 6.15: Identification of ancestral *WRR5* locus in *Eutrema salsugineum*
A - Micro-synteny analysis between the *A. thaliana* Col-0 genome (top) and *E. salsugineum* (bottom) genomes at the *WRR5* loci using CoGe SynMap2 analysis (Haug-Baltzell et al., 2017). Red blocks connected by red wedges represent syntenic regions. Genes are represented in green displaying intron and exon structure, *WRR5B* gene is shown in yellow. B – Nucleotide schematic of *A. thaliana* and *E. salsugineum* loci containing *WRR5* like proteins. Blue bubbles represent exons, scale bar represents 1 kb. C – Predicted protein architecture of *E. salsugineum* WRR5A and WRR5B like proteins using SMART prediction software (Letunic et al., 2015). Scale bar represents 100 amino acids.

The evolution of *WRR7*

Homologs of *WRR7* were found throughout the Brassicaceae genomes studied, suggesting that it evolved towards the base of the Brassicaceae family in a common ancestor to the Brassicaceae species. *WRR7* is part of the clade II DA1 family proteins, which we earlier showed to have evolved after the divergence of the Brassicaceae lineage. Therefore, we hypothesised that there would be a homolog of the ancestral NLR gene that moved into the *WRR7* position present in the closest relatives of the Brassicaceae family. It was previously hypothesised, that the *A. thaliana NRG1.3* gene (*At5G66890*) which encodes a truncated NB-ARC domain followed by a leucine rich repeat, may have split in two and the N-terminal encoding region of the ancestral *NRG1.3* gene had moved into the *WRR7* locus (Meyers et al., 2003). If this hypothesis is correct, we would expect to find an NRG 1.3-like gene that encodes the N-terminal and C-terminal ends of a traditional CC_R type NLR in the *NRG* loci of the species most closely related to the Brassicaceae family such as *Carica papaya*. To test this hypothesis, we aligned the *A. thaliana* Col-0 genome with the genome of *Carica papaya* (V5.0) using CoGE SynMap2 software and visualised the *NRG* loci using GEvo software (Haug-Baltzell et al., 2017).

We were able to identify *NRG* syntenic loci between the *A. thaliana* Col-0 genome and that of *C. papaya* (Fig 6.16). Interestingly, we did not find an intact *NRG1.3* like homolog at the *NRG* loci in *C. papaya*, in fact *NRG 1.3* seems to be entirely absent from this locus in *C. papaya* (Fig 6.16). However, we did identify the presence of an RPW8 and partial NB-ARC domain containing protein at this locus (protein ID: XP_021908385.1) that most closely resembles *WRR7* when BLASTP searched against the *A. thaliana* genome (showing 36% identity). This gene is not present in the *A. thaliana* genome (Fig 6.16) and therefore represents a prime candidate gene that could have translocated into the *WRR7* loci during the evolution of the Brassicaceae family.

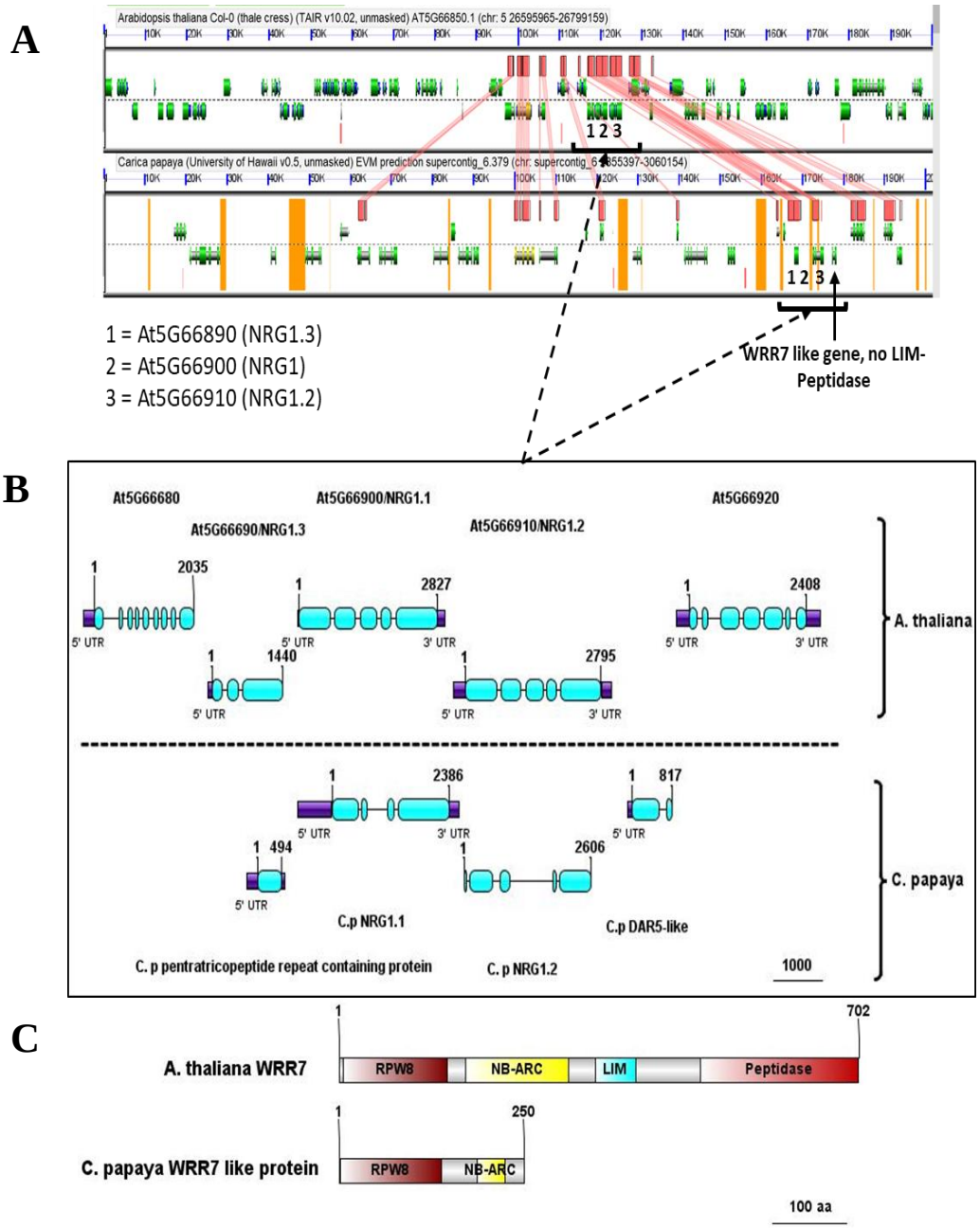


Figure 6.16: Synteny analysis of *Arabidopsis thaliana* and *Carica papaya* NRG loci

A - Micro-synteny analysis between the *A. thaliana* Col-0 genome (top) and *Carica papaya* (bottom) genomes at the *NRG* loci using CoGe SynMap2 analysis (Haug-Baltzell et al., 2017). *A. thaliana NRG* genes *NRG1.1*, *NRG1.2* and *NRG1.3* are numbered 1, 2 and 3 respectively. Red blocks connected by red wedges represent syntenic regions. Genes are represented in green displaying intron and exon structure. B- Nucleotide schematic diagram of genes located in the *NRG* loci in *A. thaliana* and *C. papaya* as viewed in J-browse software on NCBI and TAIR. Scale bar represents 1000 nucleotides, purple boxes show untranslated regions (UTR) and blue boxes indicate the exon positions in each gene. C – Protein schematic diagram of *A. thaliana* WRR7 and the WRR7 like *C. papaya* gene, protein predictions were made using SMART software (Letunic, Ivica, Doerks and Bork, 2015).

Domain swap experiments of WRR7 and DA1 family LIM-peptidase domains

To test whether the evolution of an integrated domain into the architecture of a resistance protein can be replicated, we performed domain swapping experiments using the DA1 family proteins (Fig 6.17). To perform this experiment, we replaced the LIM-Peptidase domain of WRR7 with the LIM-peptidase domains of other proteins in the DA1 family (Fig 6.17). Including the LIM-peptidase domain from DAR6 which showed close homology to that of WRR7 (Fig 6.1) as well as the LIM-peptidase domain of NLR WRR5B (DAR4). Each fusion protein was driven by the *WRR7* promoter (1993 bp upstream of the *WRR7* start codon) and fused to the *WRR7* terminator (635 bp downstream of the *WRR7* stop codon). Each fusion construct was then transformed into the *A. candida* isolate AcEM2 susceptible *A. thaliana* Col-*eds1.2-wrr7* line, which can be complemented by transformation with the Col-0 allele of *WRR7*. We hypothesised, that the most likely fusion proteins to show a response would be the *proWRR7:WRR7*$^{N\text{-}term}$*: DAR6*$^{LIM\text{-}Peptidase}$*:WRR7*^{Term} and *proWRR7:WRR7*$^{N\text{-}term}$*: WRR5B*$^{LIM\text{-}Peptidase}$*:WRR7*^{Term} because the DAR6 LIM-Peptidase is closest in homology to the LIM-Peptidase of WRR7 (Fig 6.1) and the WRR5B LIM-Peptidase domain is able to elicit immunity against *A. candida* in the WRR5 mediated immune system (Fig 3.5).

We found that none of the WRR7 fusion constructs were able confer resistance following infection by *A. candida* isolate AcEM2 in heterozygous T_1 Col-*eds1.2-wrr7* plants (Fig 6.18). However, the *proWRR7:WRR7*$^{N\text{-}term}$*:DAR3*$^{LIM\text{-}peptidase}$*:WRR7*^{Term} and the *proWRR7:WRR7*$^{N\text{-}term}$*: WRR5B*$^{LIM\text{-}Peptidase}$*:WRR7*^{Term} were able to induce a mild autoimmune phenotype (Fig 6.18). These plants displayed reduced growth, early leaf senescence and were unable to flower (Fig 6.18). The *proWRR7:WRR7*$^{N\text{-}term}$*:DA1*$^{LIM\text{-}peptidase}$*:WRR7*^{Term}, *proWRR7:WRR7*$^{N\text{-}term}$*:DAR1*$^{LIM\text{-}peptidase}$*:WRR7*^{Term} and *proWRR7:WRR7*$^{N\text{-}term}$*:DAR6*$^{LIM\text{-}peptidase}$*:WRR7*^{Term} fusion constructs showed no abnormal phenotypes compared to Col-0 plants of the same age (Fig 6.18).

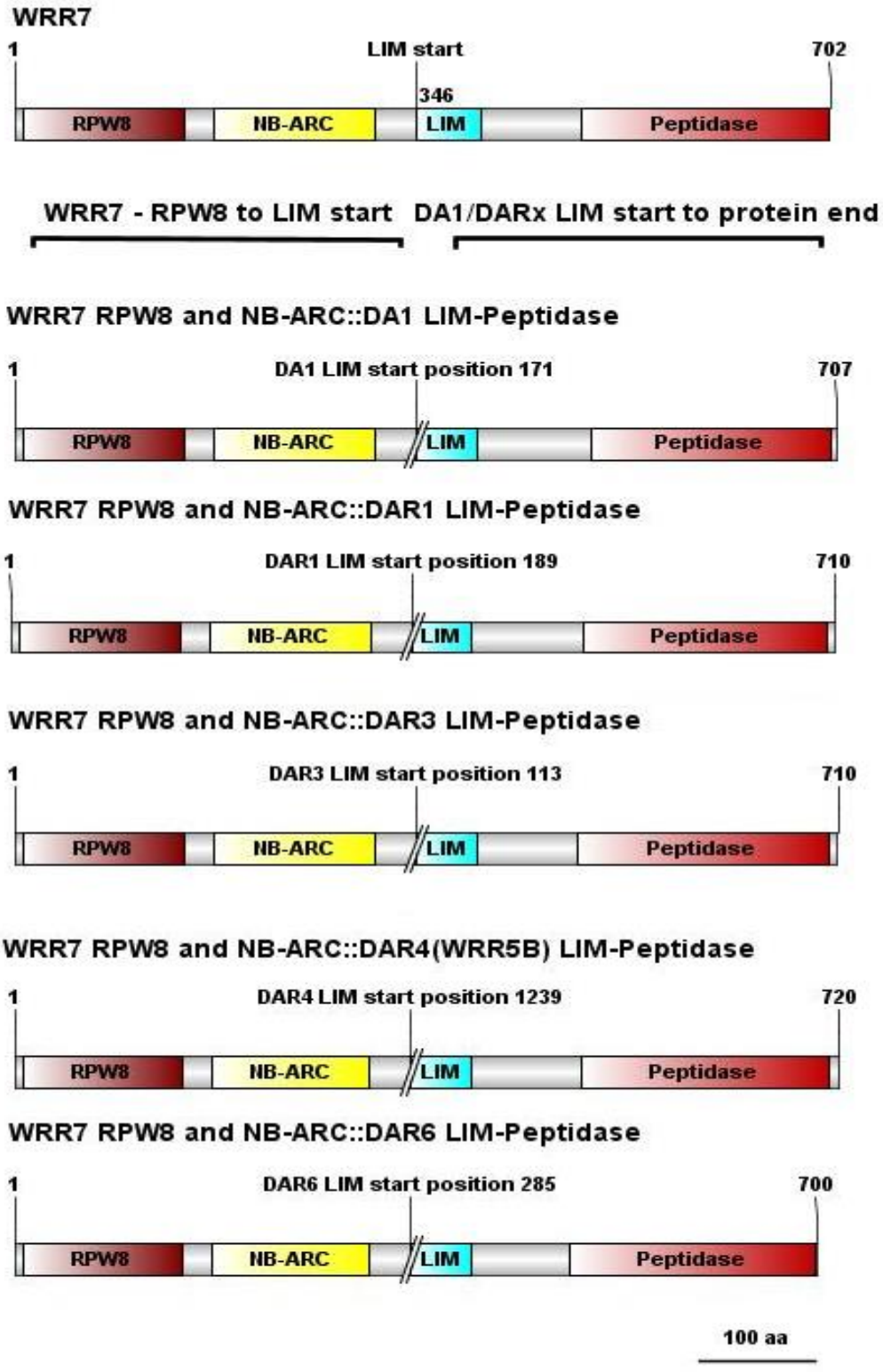


Figure 6.17: Domain swaps of the WRR7 LIM-peptidase integrated domain
Protein schematic depicting the WRR7 domain swaps performed. The WRR7 LIM-peptidase domain was swapped for the LIM-peptidase domains from DA1, DAR1, DAR3, DAR4 (WRR5B) and DAR6 proteins. Domain predictions were performed using SMART (Letunic, Ivica, Doerks and Bork, 2015). Scale bar shows 100 amino acids.

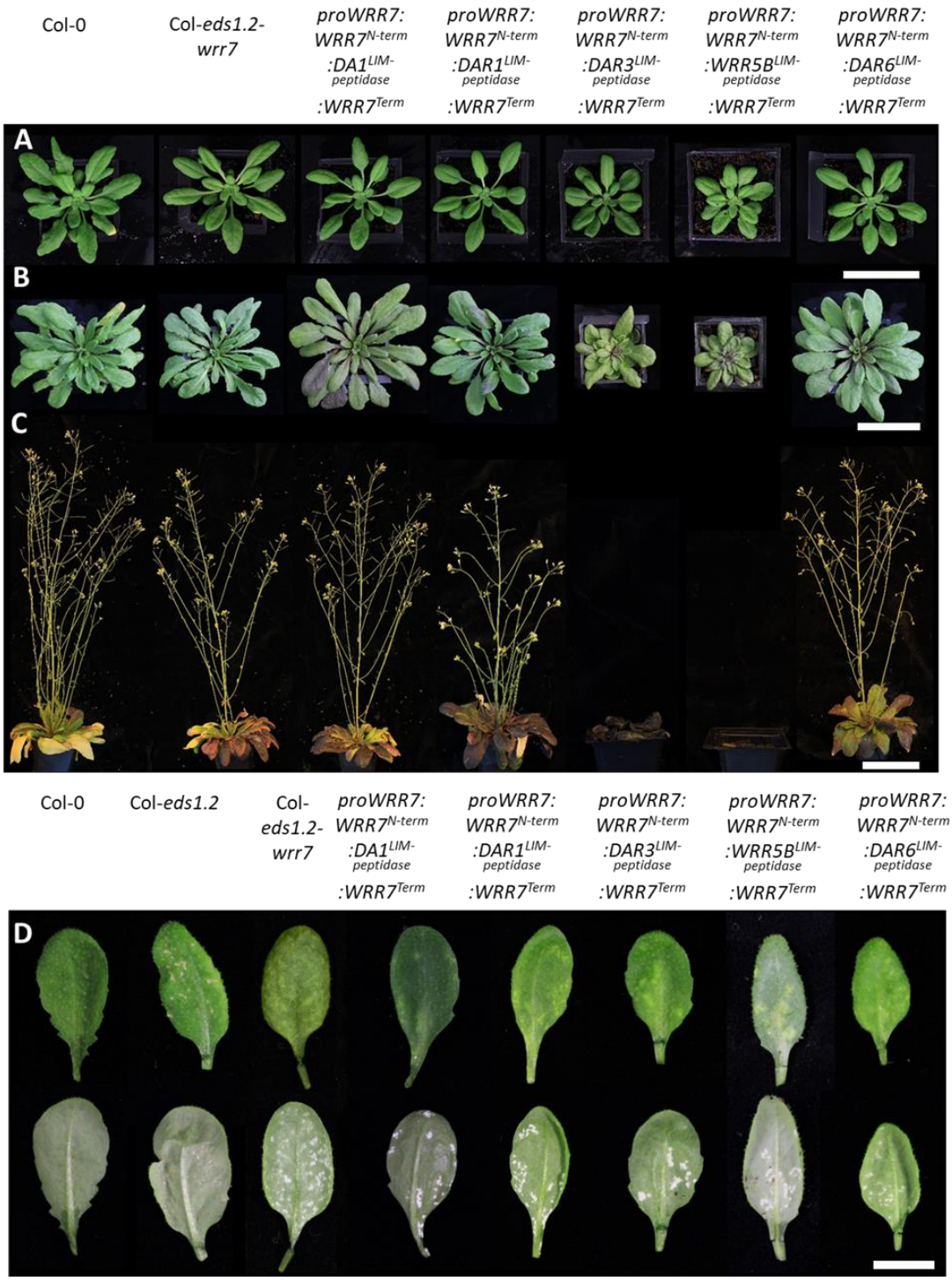


Figure 6.18: Domain swaps of DA1 family LIM-Peptidase domains onto WRR7 N-terminal region induce autoimmunity but not resistance to *Albugo candida.*

A-C Phenotype images of heterozygous T1 *Arabidopsis thaliana* Col-*eds1.2-wrr7* plants transformed with fusion constructs consisting of the *WRR7* promoter (1993 bp upstream of the start codon) with the *WRR7* sequence encoding the RPW8 and NB-ARC ($WRR7^{N\text{-}term}$) domains fused to the LIM-Peptidase encoding sequences of *DA1*, *DAR1*, *DAR3*, *WRR5B* and *DAR6* and then fused to the *WRR7* terminator region (635 bp downstream of the stop codon). A- growth after 7 weeks, B- growth after 9 weeks and C- growth after 11 weeks. D – Adult leaf images 10 days post infection phenotype images of the same lines inoculated with *Albugo candida* isolate AcEM2.

Discussion

DA1 family proteins DA1, DAR1 and DAR2 have been known to effect organ development for a number of years (Li, Yunhai et al., 2008; Peng, Yuancheng et al., 2015; Vanhaeren et al., 2017). However, the DA1 family in *A. thaliana* also contains two resistance proteins in the DA1 family (WRR5B and WRR7), as well as three other proteins DAR3, DAR6 and DAR7 that have no previously known function. We have shown that the DA1 family of proteins form two independent clades in *A. thaliana*. Clade I contains DA1, DAR1 and DAR2 that represents an ancient clade that is highly conserved in the Plantae and clade II consisting of DAR3-7 which evolved in the Brassicaceae (Fig 6.12). We were also able to show that the clade II proteins are Brassicaceae specific and likely arose through a duplication event during the evolution of the Brassicaceae. In addition, we found that overexpression of *DAR6* in *A. thaliana* ecotype Ws-2 resulted in severe phenotypic defects (Fig 6.4), suggesting that DAR6 is important in regulating the transition of several developmental programmes in *A. thaliana.* In contrast, overexpression of *DAR3* and *DAR7* caused no observable phenotypic abnormalities in Ws-2 (Fig 6.3).

Sequence analysis of clade II DA1 family LIM-peptidase protein domains (excluding resistance proteins WRR5B and WRR7), revealed that clade II proteins are evolving at an increased rate compared with clade I proteins (Fig 6.7). This suggested that clade II proteins could be acting as decoy proteins that are targeted by *A. candida* effectors, similar to the targeting of psuedokinase RKS1 by the PopP2 effector (Roux et al., 2014). Decoy proteins share structural similarity to putative pathogen effector targets within the host (Van Der Hoorn and Kamoun, 2008; Roux et al., 2014). DA1 family members contain a peptidase domain that is able to cleave E3 ligases at its active site which is characterised by the HEMMH motif (Dong, H. et al., 2017). We found that the peptidase active site motif of DAR3 has been lost and that a large insertion upstream of the DAR7 peptidase active site could have interfered with the activity of this domain (Fig 6.6), supporting the idea that these proteins could be functioning as decoys. In addition, *DAR7* is upregulated following *A. candida* infection (Fig 5.4) and has a CAMTA binding site in its promoter (Fig 6.6). *DAR7* also showed an expression profile similar to *WRR7* after *A. candida* infection and its expression profile following *A. candida* infection clusters with a group of genes that are enriched for CAMTA binding motifs in their promoters (Fig

5.4 and table 5.1), suggesting that both *WRR7* and *DAR7* could be co-regulated by CAMTA2. We also detected that DA1 family proteins become destabilised following *A. candida* infection (Fig 6.5), however this response is not consistent but this observation provides evidence that *A. candida* effectors affect these proteins following infection. Although, we speculate that DA1 family members are targeted by *A. candida* effectors and that DAR3 and DAR7 proteins could be acting as decoys, it is not possible to prove this theory until a known effector that is recognised by WRR5B or WRR7 can be identified.

Interestingly, the active site motif of the integrated peptidase domains in WRR5B and WRR7 is intact (Fig 6.6). However, neither of these domains were capable of cleaving the known target of the DA1 peptidase domains, EOD1, when fused to the N-terminal region of DA1 (Fig 6.9). Whilst the LIM-peptidase domain of DAR1 could still cleave known substrate EOD1 (Fig 6.9). Therefore, even though the active site of the WRR5B and WRR7 peptidase domains are still intact they cannot cleave the known substrate of the DA1 peptidase domain EOD1. However, the peptidase domain of WRR5B and WRR7 may have other substrates so we can not conclusively determine that these integrated decoy domains have lost their original host function.

Understanding the evolution of resistance genes and the evolutionary processes that lead to the integration of decoy domains into their architecture is important for informing future resistance gene engineering approaches. We were able to track the evolutionary origin of *WRR5A* and *WRR5B* by identifying their homologs in *Eutrema salsugineum,* where the *WRR5B* homolog does not contain DNA that encodes for the integrated LIM-Peptidase domain seen in the *A. thaliana* and *C. sativa WRR5B* alleles (Fig 6.15). Therefore, during the evolution of *WRR5B* it seems that a gene encoding a DA1 family protein translocated into the *WRR5A* and *WRR5B* locus after the divergence of the *Arabidopsis* and *Camelina* genera from the *Eutrema* genera. Subsequently resulting in a *WRR5B* allele that encoded for the integrated LIM-Peptidase domain, thus allowing *WRR5A* and *WRR5B* to confer resistance to *A. candida*.

In contrast, *WRR7* is located in a locus that contains clade II DA1 family members *DAR3*, *DAR6* and *DAR7* (Fig 6.13). A previous study suggested that the N-terminal region of an ancestral *NRG1.3* gene, encoding a protein with a RPW8 domain and a

partial NB-ARC domain had translocated into the locus containing *WRR7* in *A. thaliana* (Meyers et al., 2003). We could not identify an intact ancestral CC_R homolog of *NRG1.3* in *Carica papaya*, the most closely related genome that we could study to the divergence of the Brassicaceae that would not contain clade II DA1 family proteins. However, we did find a RPW8 and partial NB-ARC domain containing protein at the *NRG* locus in *C. papaya* that most closely resembles *WRR7* when BLASTP searched against the *A. thaliana* genome (Fig 6.13). Therefore, we find it likely that, in the case of *WRR7,* an ancestral form of the *C. papaya* RPW8 and NB-ARC domain containing protein moved into the *WRR7* locus and fused with a DA1 family gene, that probably arose from a duplication of *DAR6.*

Therefore, we have identified two scenarios leading to the integration of a LIM-peptidase domain into an NLR during the course of the Brassicaceae family evolution. One scenario where the most probable integration event involves the translocation of a gene encoding a truncated NLR into an integrated domain encoding gene, forming the LIM-Peptidase encoding resistance gene *WRR7*. The second scenario most likely involves the translocation of a gene encoding the LIM-Peptidase integrated domain into the locus containing a gene encoding a canonical NLR, subsequently resulting in the evolution of the *A. thaliana WRR5B* allele. Therefore, both WRR5B and WRR7 gained their LIM-peptidase domains via different processes proceeding the emergence of the clade II DA1 family proteins in the Brassicaceae. WRR7 at some point close to this event and WRR5B much later on just before the divergence of the *Camelina* and *Arabidopsis* genera but after the divergence of *Eutrema*.

We then attempted to simulate the evolution of *WRR7* using domain swapping experiments, fusing LIM-Peptidase domains from other DA1 family members onto the RPW8 and NB-ARC domain of WRR7. We found that none of the fusion proteins generated could provide resistance against *A. candida* isolate AcEM2 (Fig 6.18). Therefore, the simple fusion of NLR to integrated domain was not enough to be able to elicit an immune response. We did manage to induce a mild autoimmune response by fusing the LIM-peptidase domains of DAR3 and WRR5B onto the WRR7 N-terminal region (Fig 6.18). This may provide a route to the evolution of functional NLRs with integrated domains through mild immune activation that could provide enough resistance to give a selective advantage to plants displaying this

phenotype. The mild activation of autoimmunity could then facilitate host-pathogen co-evolution, resulting in more highly specific and attuned resistance genes, able to activate resistance only when in the presence of the target phytopathogen.

The integration of the LIM-Peptidase integrated decoy domain into the architecture of both WRR5B and WRR7 occurred within the Brassicaceae lineage in two different events once at the base of the family and once prior to the divergence of the *Camalina* and *Arabidopsis* genera. However, there are LIM-peptidase encoding resistance genes of both the TNL and CC_R classes outside of the Brassicaceae family (Srivastava and Verma, 2015; Sarris, P. F. et al., 2016). Therefore, the integration of the LM-Peptidase into a plant NLR has happened on at least four separate occasions (Fig 6.12). We have shown that *WRR5B* and *WRR7* both confer resistance to Brassicaceae white rust pathogen *A. candida* (Fig 3.5 and 3.10). However, the presence of LIM-peptidase domain containing resistance genes occurring in plant families that are not targeted by *A. candida,* suggests that the non-Brassicaceae LIM-peptidase encoding resistance genes potentially recognise other phytopathogens. If this is the case, then the LIM-peptidase domain offers an exciting prospect for resistance gene engineering because multiple phytopathogens will target this domain.

Chapter 7

General discussion

Identification of novel *White Rust Resistance* genes

In this thesis, I provide evidence that the resistance genes *WRR5A, WRR5B* and *WRR7* from the *Arabidopsis thaliana* Col-0 genome all confer resistance to *A. thaliana* infecting phytopathogen *Albugo candida* race 4 isolate AcEM2. The *WRR5A* & *WRR5B* and *WRR7* resistance mechanisms operate independently of one another and are independent of the previously identified Col-0 white rust resistance gene, *WRR4A* (Borhan, M. Hossein et al., 2008). *WRR5A* and *WRR5B* are both required for the *WRR5* resistance mechanism. Therefore, the Col-0 genome contains four resistance genes and three resistance gene mechanisms that confer resistance to AcEM2.

WRR5A and WRR5B operate by the sensor-helper model of NLR activation

WRR5A and *WRR5B* encode a pair of TNLs that are arranged in a tandem head to head orientation in the *A. thaliana* genome and share a promoter region (Xu et al., 2015). The presence of both of these genes is necessary to stimulate an autoimmune response conferred by the *WRR5B-C1340Y* autoimmune allele (also referred to as *chs3-2D)* (Xu et al., 2015). We have shown that *WRR5A* and *WRR5B* are both required to cause an immune response to AcEM2 in *A. thaliana.* We also found that overexpression of *WRR5A* in the absence of *WRR5B* led to an autoimmune response in *A. thaliana* plants and that when *WRR5A* and *WRR5B* were both expressed, this phenotype was supressed. Furthermore, we demonstrated that WRR5A and WRR5B proteins form a heterodimeric complex in *A. thaliana*. Therefore, we propose a model whereby WRR5B inhibits the autoactivation of WRR5A in non-infected cells, similar to the RGA4 and RGA5 model where RGA5 inhibits the auto activity of RGA4 (Césari et al., 2014). The traditional sensor-helper model of NLR activation suggests that only the helper NLR needs to perform traditional immune signalling governed by the exchange of ADP for ATP to cause immunity (Cesari et al., 2014; Sohn et al., 2014; Wang, J., Wang, et al., 2019). However, we found that the

ADP/ATP binding P-loop region of both WRR5A and WRR5B are required to stimulate the autoimmune response elicited by the WRR5B-C1340Y mutation. Therefore, the mechanism employed by WRR5A and WRR5B exhibits similarities to previously identified sensor-helper NLR mechanisms, such as RPS4/RRS1 and RGA4/RGA5 but also shows a distinct difference in the requirement for functional P-loop motifs in both NLRs (Césari et al., 2014; Sohn et al., 2014). Exactly how such a dual P-loop mechanism would operate is still unknown. However, a possible explanation is that the WRR5B LIM-peptidase domain interferes with the release of WRR5A from the WRR5A-WRR5B heterodimer until its P-loop changes in conformation following the exchange of ADP for ATP and that the WRR5A P-loop is required for downstream immune signalling and the formation of a reistomome like structure (Wang, J., Hu, et al., 2019).

WRR5A-YFP and WRR5B-YFP both localised to the plasma membrane in transient expression experiments performed in *Nicotiana benthamiana* and WRR5A also exhibited nuclear localisation. Cellular fractionation experiments of constituently overexpressed WRR5A and WRR5B in *A. thaliana* also showed WRR5A and WRR5B to be present in the microsomal and nuclear fractions in non-infected plants and that WRR5A was enriched in the nuclear fraction after infection with *A. candida* isolate AcEM2. Therefore, we propose that after infection with AcEM2, WRR5A and potentially WRR5B relocates from the membrane to the nucleus where it stimulates defence activation. However, we have not been able to determine whether WRR5A and WRR5B move as a complex or whether WRR5A translocates to the nucleus independently of WRR5B.

Previous studies have attempted to find other proteins involved in the WRR5A and WRR5B-C1340Y autoimmune response. One approach taken, was to look for suppressors of the autoimmune response using an EMS screens. One screen revealed that, mutations in the known TNL signalling components SAG101 and EDS1 as well as in WRR5A could supress the autoimmunity induced by the WRR5B-C1340Y autoimmune mutant and that *pad4-1* mutants could not fully rescue the autoimmune phenotype (Xu et al., 2015). A different screen identified that mutants of *Indole-3-Butyric Acid Response 5* (*ibr5*) could suppress autoimmune responses induced by the *chs3-1* mutants and that IBR5 can interact with the WRR5B TIR domain in a yeast-2-hybrid study as well as HSP90 (Liu, J. et al., 2015). In addition, a screen of

chemicals that could inhibit the WRR5B-C1340Y autoimmune response found that, Ro- 8-4304 could suppress WRR5B-C1340Y autoimmunity by targeting the methylosome complex (Huang et al., 2016). Impairing the spliceosome machinery, resulting in alternative transcripts encoding truncated WRR5B at the LIM domain (Huang et al., 2016). More recently it has been shown that NRG and ADR1 proteins are involved in the downstream signalling of TNLs including the autoimmune response governed by WRR5A and WRR5B-*chs3*-2D (Castel et al., 2018; Wu, Z. et al., 2018). We were able to confirm that this observed autoimmune response in transient expression experiments required NRG1, however crosses between the CW5 RIL with a T-DNA knock out line of all three *nrg* genes yielded no susceptible mutants and surprisingly crosses between CW5 and the Col-*adr1* triple knock out did yield susceptible F_2 individuals. This suggests that the resistance mechanism mediated by *WRR5A* and *WRR5B* to *A. candida* may operate independently of the autoimmune responses observed in transient expression experiments with the autoimmune allele of *WRR5B.*

To further elucidate any proteins that are involved with the WRR5A and WRR5B mechanism, we performed IP-MS analysis of our *A. thaliana* Ws-2 lines that were overexpressing *WRR5A-V5* and *WRR5B-HF*. Several potential interacting proteins were identified by IP-MS, these included proteins involved in post-translational protein processing and folding such as heat-shock proteins and chaperones. Post-translational processing of NLR proteins is a well-established mechanism of regulating immune responses, for example RPM1 is a regulated by HSP90 in co-ordination with co-chaperones RAR1 and SGT1b (Hubert et al., 2003; Holt, Belkhadir and Dangl, 2005). We identify through IP-MS that both WRR5A and WRR5B are associated with HSP70 and co-chaperone ATJ3, confirming the earlier finding that these two proteins are post-translationally regulated by HSPs (Liu, J. et al., 2015). As well as proteins involved in post-translational modification, we identified the plasma membrane aquaporin protein PIP2 as interacting with both WRR5A and WRR5B. PIP2 is often used as a plasma membrane marker protein, further backing up our finding that the WRR5 complex localises to the plasma membrane in uninfected plants (Czech, 2000). In both IP experiments, we identified lipoxygenase 2 (LOX2) as the protein co-immunoprecipitating in the greatest abundance with WRR5A and WRR5B other than WRR5A and WRR5B themselves.

LOX2 is predominantly located in the chloroplast but relocates to the plasma membrane under elevated Ca^{2+} conditions (Walther, Wiesner and Kuhn, 2004; Järving et al., 2012; Mochizuki and Matsui, 2018). Therefore, this interaction provides a potential mechanism for the WRR5 complex to recognise the presence of *A. candida,* as infection by phytopathogens is known to trigger changes in cellular Ca^{2+} levels (Zipfel and Oldroyd, 2017). As well as interacting proteins that are associated with the plasma membrane, we also identified nuclear associated transcriptional regulator proteins BRM and NOT3 as well as histone H4 in the WRR5A IP experiment. None of these proteins were identified in the WRR5B IP samples backing up our confocal localisation findings that WRR5A has nuclear localisation as well as plasma membrane localisation but that WRR5B is more exclusive to the plasma membrane under non-infected conditions. The interaction of WRR5A with transcriptional regulators that act in remodelling chromatin, provides a link between WRR5 immunity and downstream defence activation (Li, C. et al., 2016).

A non-canonical CC_R type resistance protein WRR7 provides resistance against *Albugo candida*

WRR7 is an atypical resistance gene that belongs to the CC_R class due to the presence of an RPW8 domain at its N-terminus. It is unusual as it contains only a partial NB-ARC domain with a non-functional P-loop and has no leucine rich repeat, in addition it contains an integrated LIM-peptidase domain at its C-terminus. We demonstrated that *WRR7* was responsible for the resistance response observed in Col-*eds1.2* plants by generating Col-*eds1.2-wrr7* mutant lines and showing that the resistant cell death response was abolished in these plants. We then complemented these lines with *gWRR7* and showed that the necrotic resistance phenotype observed in both Col-*eds1.2*, could be recovered by the insertion of *gWRR7* into the Col-*eds1.2-wrr7* background. We also attempted to generate a resistance response in AcEM2 susceptible ecotype Ws-2, however we found that transformation of Ws-2 with *gWRR7* was unable to fully confer a resistance response. Ws-2 contains a copy of *WRR7* in its genome which contains only one non-synonymous point mutation at position 528 in the amino acid sequence. There are two possible scenarios why the *gWRR7* Col-0 allele didn't fully complement the susceptible Ws-2 phenotype, either

another gene that is required for *WRR7* immunity is non-functional in Ws-2 or that the non-functional allele of Ws-2 interferes with the Col-0 *WRR7* allele for effector recognition and therefore the threshold for activation is not reached resulting in partially susceptible plants.

We found that *WRR7* mediated immunity requires upregulation of *WRR7* transcription proceeding infection to activate immunity. Therefore, transcriptional regulators are key components of the *WRR7* resistance mechanism. Our forward genetic screen identified a transcription factor (*CAMTA2*), a chromatin remodelling complex (*CHR4*) and a spliceosome associated protein (*MAC7*) that all regulate the transcription of *WRR7*.

Regulation of transcription is highly important in other resistance gene systems. For example, *RPP8* is upregulated in *A. thaliana* after infection by *Hyaloperonospora arabidopsidis* and the transcripts of TNL *SNC1* are highly regulated by multiple *modifier of SNC1* (*MOS*) genes, including binding of repressor proteins to regulatory regions, alteration of chromatin structure by changes to the methylation pattern of *WRR7* and regulation of transcript splicing (Li, Yongqing et al., 2007; Mohr et al., 2010; Xu et al., 2012a; Xia, S. et al., 2013; Lai and Eulgem, 2018).

The transcription of *WRR7* is activated by the binding of calcium responsive transcription factor CAMTA2 to the *WRR7* promoter element. CAMTA2 has previously been reported to be functionally redundant with two other transcription factors CAMTA1 and CAMTA3 and these transcription factors are known to be involved with the regulation of defence associated genes such as *EDS1* (Kim, Y. et al., 2013; Jacob et al., 2017; Yuan, Du and Poovaiah, 2018; Sun et al., 2020). There is debate in the literature as to whether CAMTA transcription factors activate or suppress plant defences, with recent studies showing CAMTA3 as having a repressive function in plant immunity (Lolle et al., 2017; Sun et al., 2020). However, we have shown that in the case of *WRR7* resistance, CAMTA2 acts as a transcriptional activator of *WRR7*. In addition, we found that CAMTA2 acts independently of CAMTA1 and CAMTA3 in regulating *WRR7* transcription. This activation is highly specific and indicates that the recognition of AcEM2 by *A. thaliana* requires a calcium signalling event. Exactly what the calcium signal is that AcEM2 induces upon infection is yet to be determined. However, we find it likely

that calmodulin proteins CaM2 or CaM3 are likely to decode this signal and interact with CAMTA2, stimulating *WRR7* expression. Interestingly, the CNGC19 channel was upregulated following *A. candida* infection and could represent a Ca^{2+} channel that could be responsive to *A. candida*. Another interesting question that arises around this signalling system is, how the Ca^{2+} signal elicited by *A. candida* specifically involves CAMTA2 and not the highly similar CAMTA1 and CAMTA3 proteins in activating *WRR7* transcription.

WRR7 transcript levels are modulated not only by CAMTA2 but also by chromatin remodelling complex CHR4. CHR4 is important for the epigenetic regulation of genes and has been shown to specifically regulate the trimethylation state of both histone 3 lysine 27 (H3K27me3) and histone 3 lysine 4 (H3K4me3) in *Oryza sativa* and in *A. thaliana* (Hu et al., 2012; Sang et al., 2020). H3K27me3 marks are associated with repression of genes whereas H3K4me3 marks are associated with gene activation as well as being correlated with the histone 2 variant H2A.Z which alters the nucleosome structure and enhances gene activation (Hu, G. et al., 2013; Chen, X. et al., 2015; Jambhekar, Dhall and Shi, 2019). Intriguingly, the *CHR4* T-DNA knock out mutant *chr4-2* showed bivalent hypermethylation of both H3K27me3 and H3K4me3 marks around the *WRR7* loci in comparison to Col-0 control plants and the *WRR7* promoter is enriched for the H2A.Z histone variant (Zander et al., 2019; Sang et al., 2020). The hypermethylation of the *WRR7* loci with both H3K23me3 and H3K4me3 in the *chr4*-2 mutant is unusual as both of these marks act in an antagonistic manner to one another. However, bivalent H3K27me3 and H3K4me3 marks are associated with repressed genes that are primed for expression by the removal of one of the marks, in plants this bivalent mark is associated with genes induced after cold stress (Bernstein et al., 2006; Jambhekar, Dhall and Shi, 2019; Zeng et al., 2019). Therefore, *WRR7* is held in a poised state by the presence of H3K27me3 and H3K4me3 bivalent mark, and that removal of H3K27me3 is important for the transcription of *WRR7* after *A. candida* infection. Our CHR4, P23S mutation was located upstream of the plant homeodomain (PHD) which reads the methylation state of H3 N-terminal marks and is particularly associated with the H3K4me3 mark (Sanchez and Zhou, 2011). PHD containing proteins are known to specifically interact with the H3K4me3 mark through their PHDs and tandem chromodomains and recruit co-activators that stimulate gene

expression (Sims et al., 2005; Hyun et al., 2017). Proline residues residing on flexible loops upstream of the PHD have been found to occupy the histone binding site of the PHDs in the absence of ligands and are believed to act as an autoregulatory mechanism (Ramón-Maiques et al., 2007; Li, Yuanyuan and Li, 2012). Therefore, the EMS 34 mutation of P23S in CHR4 may prevent this auto-inhibitory mechanism of the PHD. It is conceivable that this mechanism is the switch that regulates binding of CHR4 to H3K4me3 residues in the presence of *A. candida* and that without this P23S mutation prevents the binding of CHR4 to the *WRR7* regulatory regions resulting in reduced expression of *WRR7* following infection. Alternatively, this mutation could affect the localisation of this protein because the P23S mutation resides in a nuclear localisation signal that runs from amino acid positions 7-35 of the CHR4 protein sequence (Kosugi S. et al., 2009).

In addition, co-immunoprecipitation-mass spectrometry experiments using CHR4 as the bait identified CAMTA3 as an interacting factor in *A. thaliana* inflorescences revealing that Ca^{2+} responsive transcription factors can interact with CHR4 (Sang et al., 2020). Therefore, it is likely that CHR4 is important for the removal and regulation of one of the H3K27me3 or H3K4me3 epigenetic marks, most likely the H3K27me3 mark, leading to the transcription of *WRR7* via the binding of CAMTA2 to the *WRR7* promoter. Exactly how CAMTA2 and CHR4 act in concert to activate *WRR7* mediated immunity is unknown. Studying whether CAMTA2 physically interacts with CHR4 during an *A. candida* triggered immune response may provide insights into the mechanistic interaction of transcription factors with chromatin remodelling complexes and how these two key components of transcriptional regulation function together to bring about highly specific responses in plants.

Recently it has been shown that transcription and splicing occur simultaneously in plants (Li, S. et al., 2020) and after transcripts are produced, they have to be correctly spliced before being exported from the nucleus. These processes provide additional distinct regulatory phases where resistance gene transcripts can be regulated. We identified a spliceosome associated protein (MAC7) as being important for the correct splicing of *WRR7* transcripts. MAC7 is a conserved RNA-helicase protein found across the kingdoms of life, in plants it plays an important role in plant defence as *mac* mutants have downregulated defence genes (Jia et al., 2017). The MAC complex plays a key role in recruiting the splicing machinery to

the chromatin (Li, S. et al., 2020) and MAC7 itself has been shown to be involved in pre-mRNA processing as well as being important in the production of miRNA and siRNAs (Jia et al., 2017). Exactly how MAC7 is involved in the regulation of *WRR7* transcripts remains to be determined, although a model whereby MAC7 is involved in prolonging the half-life of *WRR7* transcripts seems the most likely as in the EMS 146 MAC7 mutant line *WRR7* activation is enough to cause cell death but not enough to completely inhibit pathogen growth.

Another mutant line that showed a similar phenotype to the MAC7 mutant line was EMS 1, which had a mutation in mitogen activated protein kinase kinase kinase gene *MAP3Kδ4*. This mutant line exhibited extensive cell death that was not enough to fully block pathogen growth, with a small number of pustules developing ten days post infection. MAP Kinases are important regulators of plant immunity and form part of the downstream signalling networks of both TNLs and CNLs and as such are targeted by pathogen effectors, a MAP3K protein, MAPKKKε is a known target of Oomycete phytopathogen *Phytophthora infestans* (King et al., 2014; Bi, G. and Zhou, 2017). MAP3Kδ4 has a known function in regulating ABA signalling (Lumba et al., 2014). ABA is a phytohormone that acts antagonistically to plant defence hormone SA (Meguro and Sato, 2014; Kusajima et al., 2017). Therefore, ABA signalling needs to be supressed during a salicylic acid induced immune response. Identification of an AcEM2 susceptible line with a mutation in *MAP3Kδ4* suggests that this MAP3 Kinase plays a role in the suppression of ABA signalling during the *WRR7* mediated immune response, whether this protein is involved more generally as a suppressor of ABA during plant immune responses is yet to be determined.

Therefore, our forward genetic screen implicates Ca^{2+} signalling, chromatin remodelling, and abscisic acid signalling as being important components of *WRR7* mediated immunity.

Calcium signalling and the phytobiome

Plants reside in an ecological niche that is perpetually in contact with a community of other organisms, this community is collectively known as the phytobiome. In order to survive, plants have to be able to perceive which organisms inhabit the phytobiome and in some cases interact with other organisms to either recruit

symbiotic partners or to defend against invading organisms (Leach et al., 2017; Zipfel and Oldroyd, 2017). Distinguishing between commensal, symbiotic or pathogenic organisms is crucial for the survival of plants. One common event triggered by the interaction of the phytobiome with a host plant, is the induction of changes in Ca^{2+} concentrations. Changes in cytosolic Ca^{2+} concentrations are induced by damage or disruption to plant cellular surfaces, either through macroscopic damage caused by herbivory or through microscopic disruption to cellular structures by the invasion of pathogenic or symbiotic microorganisms (Ranf et al., 2011; Vadassery et al., 2012). Perception of different groups of organisms operates through overlapping signalling networks, for example CERK1 a co-receptor required for the recognition of chitin in *A. thaliana* to activate MTI is also requited to recognise arbuscular mycorrhizal fungi in *O. sativa* (Miya et al., 2007; Carotenuto et al., 2017; MacLean, Bravo and Harrison, 2017). In addition, other MTI associated proteins such as receptor-like kinases BIK1 additively enhance cytosolic Ca^{2+} influxes following infection (Li, L. et al., 2014). Therefore, cell surface receptors are highly important for recognising micro-organisms that are both mutualistic and pathogenic resulting in Cytosolic Ca^{2+} influxes across the plasma membrane.

Ca^{2+} signatures are diverse in plant cells due to the development of discrete calcium sources and sinks that are separated by membranes. Ca^{2+} concentrations can be rapidly altered in discrete sub-cellular compartments, such as the cytosol, nucleus, plastids and the vacuole by the passage of Ca^{2+} through Ca^{2+} channels in membranes (Kudla, Batistič and Hashimoto, 2010; Duszyn et al., 2019). In addition, Ca^{2+} signatures differ in amplitude, frequency and duration resulting in a fine-tuned stimulus response system that operates as a master-regulator of plant responses (Kudla, Batistič and Hashimoto, 2010). Ca^{2+} signatures are decoded by calcium decoding proteins that bind cellular calcium via EF-hand domains, culminating in the activation of the discrete signalling networks leading to the induction of different developmental pathways (Dodd, Kudla and Sanders, 2010; Kudla et al., 2018; La Verde, Dominici and Astegno, 2018). The intricacies of the Ca^{2+} signalling system are beginning to be understood. Symbiotic relationships between plants with Rhizobial bacteria and Arbuscular mycorrhizal fungi are triggered by specific nuclear Ca^{2+} spiking events that are highly similar but have different frequencies, leading to distinct symbiotic responses in the plant (Kosuta et al., 2008; Genre et al.,

2013). On the other hand, MTI can be upregulated by rapid cytosolic Ca^{2+} influxes across the plasma membrane and subsequent membrane depolarisation causing Ca^{2+} oscillations for up to 30 minutes (Ranf et al., 2011; Li, L. et al., 2014; Keinath et al., 2015). ETI responses require a more prolonged increase in cellular Ca^{2+} (Grant et al., 2000), and the production of SA, a key phytohormone integral to ETI activation, is controlled by Ca^{2+} responsive transcription factor CBP60g which regulates genes such as *ICS1* and *EDS5* which are crucial for the biogenesis and transport of SA in the chloroplast (Wang, L. et al., 2009; Ding and Redkar, 2018).

Recently, Cyclic Nucleotide-Gated ion Channels (CNGC) channels that transport Ca^{2+} across membranes have been implicated in facilitating the influx of Ca^{2+} into the cytosol and nucleus during MTI, ETI and symbiotic interactions (Charpentier et al., 2016; Duszyn et al., 2019; Meena et al., 2019; Tian, W. et al., 2019; Yu, X. et al., 2019). CNGC channels physically associate with CaM proteins that can regulate the resulting Ca^{2+} oscillations by Ca^{2+}-CaM mediated blocking of the associated channel (Pan et al., 2019). Following the induction of specific Ca^{2+} signals, Ca^{2+} decoding proteins such as CaMs or Calcium-dependant protein kinases (CPKs) then go on to upregulate specific regulatory pathways resulting in physiological responses to various stimuli. For example, CPK4/5/6 and 11 upregulate WRKY transcription factors following Ca^{2+} dependent ETI responses resulting in immunity associated gene activation (Gao, X. et al., 2013). Our finding that *WRR7* is activated by CAMTA2, a CaM binding transcription factor suggests that a similar mechanism is utilised in an *A. thaliana* immune response to *A. candida*. Whereby, *A. candida* infection disrupts the *A. thaliana* cellular surface stimulating a Ca^{2+} influx through a plasma membrane localised CNGC channel, this influx is then decoded, likely by CaM2, that binds to CAMTA2 which goes on to stimulate *WRR7* expression resulting in a cell death response.

Decoding complex Ca^{2+} signatures during plant-microbe interactions is a highly complex area of study. Understanding how plants utilise Ca^{2+} to distinguish between mutualistic, pathogenic or commensal micro-organisms and how this differs to their responses to other Ca^{2+} regulated abiotic stresses or developmental stimuli is integral to understanding plant responses in general. We have a tendency to separate plant-microbe interactions into discrete pathways that are isolated into ‘pathogenic’ or ‘symbiotic’ interactions and although the downstream signalling leads to the

upregulation of diverse developmental pathways, the initial perception of pathogenic and symbiotic interactions is starkly similar and should be seen as a more general interaction event, irrelevant of whether the micro-organism is pathogenic, mutualistic or commensal. In a similar manner, both ETI and MTI are intrinsically linked by Ca^{2+} disruption and without MTI stimulated Ca^{2+} influxes then ETI would not be able to operate. To this end, it is increasingly apparent that the 'two layers' of plant immunity are much more closely linked and that the signalling of one is mutually dependent on signalling events from the other.

The finding that *WRR7* is another defence gene that is reliant on Ca^{2+} signalling to trigger its response offers another useful system to study in decoding the Ca^{2+} response and figuring out how a plant can activate a specific response against a particular phytopathogen. Only by unpicking the details of more pathways that are responsive to Ca^{2+}, like the *WRR7* signalling pathway, will we be able to fully appreciate the role of Ca^{2+} signalling in regulating plant responses to organisms in the phytobiome.

Chromatin remodelling and the epigenetic regulation of plant immunity

Genetic pathways encoding proteins that are important for the production of key metabolites in eukaryotes can be co-regulated if they fall into gene clusters (Nützmann and Osbourn, 2015; Yu, N. et al., 2016). Co-regulation is governed by key chromatin marks, such as H3K27me3 and H3K4me3 and the deposition of histone variants e.g. H2A.Z that can be altered by chromatin remodelling complexes (Yu, N. et al., 2016; Ojolo et al., 2018). Plant resistance genes can also form gene clusters such as the *RPP5* cluster which contains seven NLRs including *SNC1, RPP4* and *RPP5* (Meyers et al., 2003; Yi and Richards, 2007). Resistance gene clusters, like their metabolite counterparts, can be co-regulated and chromatin remodelling factors are increasingly being shown to be involved in the regulation of plant immunity, for example CHR5 positively regulates the *RPP5* cluster (Zou et al., 2017). During our experiments we identified two chromatin remodelling complexes as being implicated in *A. thaliana* immune responses to *A. candida*. Firstly, we found that BRAHMA (BRM) interacted with WRR5A by Co-IP and secondly, we

identified CHR4 as being important for regulating *WRR7* immunity as part of a forward genetic screen.

In *A. thaliana* there are 41 chromatin remodelling complexes that fall into the Sucrose non-fermentable 2 (SNF2) family of proteins that regulate chromatin structure and marks based on ATP hydrolysis (Knizewski, Ginalski and Jerzmanowski, 2008). CHR4 is part of the Mi-2 subfamily which also contains Pickle (PKL) and is closely related to CHR5, a homolog of CHD1 in *A. thaliana*, whereas BRM falls into the SNF2 related subfamily (Knizewski, Ginalski and Jerzmanowski, 2008). The finding that WRR5A interacts with BRM and that *WRR7* mediated immunity requires CHR4 adds to the growing body of literature linking resistance genes with chromatin regulation. The most heavily studied NLR in relation to chromatin remodelling complexes in plants is *SNC1* that is regulated not only by CHR5 but also by chromatin remodelling complexes Splayed (SYD) and Decrease in DNA methylation 1 (DDM1) as well, although exactly how all these interactions play out is still unknown (Ramirez-Prado et al., 2018). Another chromatin remodelling complex that is known to regulate immune associated genes is Photoperiod independent flowering 1 (PIE1) that controls the deposition of gene activating histone H3 variant H2A.Z and acts in co-ordination with PKL to regulate repressive histone mark H3K27me3 deposition (March-Díaz et al., 2008; Carter et al., 2018).

Therefore, there is a growing body of literature implicating chromatin remodelling complexes with regulation of resistance genes and we provide evidence that chromatin remodelling complexes are important in the regulation of the WRR5 and WRR7 immune pathways. We speculate that these complexes could have wider functional importance in regulating clusters of resistance genes following pathogen invasion. It is interesting to note that *chr4-2* mutant plants are hypermethylated for both H3K4me3 and H3K27me3 marks as well as being enriched for H2A.Z (Zander et al., 2019; Sang et al., 2020). Therefore, it is likely that CHR4 plays a role in releasing the H3K27me3 mark during pathogen invasion facilitating *WRR7* gene expression due to the activating H3K4me3 and H2A.Z marks. Further exploration of the role that chromatin remodelling complexes play in regulating resistance gene expression is needed and the WRR7 mechanism provides a particularly interesting model for further exploring this interaction.

ABA signalling in plant immunity

ABA is an important phytohormone that regulates plant developmental pathways, particularly in relation to abiotic stress responses (Finkelstein, 2013; Sah, Reddy and Li, 2016; Chen, K. et al., 2020) . It also has several roles in plant-microbe interactions where it acts antagonistically to SA, acting as an immune suppressor as well as aiding in the formation of mycorrhizal associations (Jiang et al., 2010; Liu, S. et al., 2015; Martín-Rodríguez et al., 2016). ABAs role as a negative regulator of ETI means that it needs to be downregulated during an immune response and as such is an active target for pathogens, which attempt to upregulate the ABA pathway during infection (de Torres-Zabala et al., 2007; Lievens et al., 2017; Peng, Z. et al., 2019). Exactly how ABA is repressed or maintained at low concentrations during infection is still unclear, although a recent study showed that genes involved in the production of ABA were downregulated during early infection of root pathogen *Verticillium longisporum* (Behrens et al., 2019). The core ABA signalling network involves the perception of ABA by Pyrabactin resistance 1 and PYR1-like receptors that upregulate SnRK2s that phosphorylate ABA-responsive element binding factors (ABFs) that result in the regulation of downstream developmental or response pathways (Lumba et al., 2014; Chen, K. et al., 2020). The regulation of SnRK2s is a crucial step in controlling ABA responses, SnRK2s are activated by phosphorylation of MAP3Ks and repressed by Protein phosphatase 2C (PP2C) mediated dephosphorylation (Lumba et al., 2014; Chen, K. et al., 2020). We identified MAP3Kδ4 as playing a role in *WRR7* mediated immunity and that when this protein was mutated the hypersensitive response elicited by ETI was not strong enough to fully inhibit *A. candida* growth, resulting in the formation of pustules 10 dpi. MAP3Kδ4 has previously been identified as a negative regulator of ABA signalling due to its interaction with PP2Cs and ABFs (Lumba et al., 2014). We also found that MAP3Kδ4 was a target of *Ralstonia psuudosolanacearum* effectors. Therefore, we propose that during *WRR7* mediated immunity MAP3Kδ4 plays an active role in repressing ABA signalling, potentially through the activation of PP2Cs that dephosphorylate SnRK2s. We also speculate that MAP3Kδ4 may be involved in broader ETI responses in regulating ABA signalling during plant immune responses and is also an important target for pathogens to manipulate ABA responses.

The LIM-Peptidase protein family and their role in plant immunity

The LIM domain containing proteins have been sub-divided into four distinct classes, one of which is plant specific, and is associated with the zinc metallopeptidase domain (Pfam: PF12315) (Zhao, M. et al., 2014). The first identified member of this family was DA1 and it is therefore known as the DA1 family, both WRR5B and WRR7 are members of this protein family. Interestingly, the DA1 family has undergone multiple duplication events during the evolution of plant species, at least once in monocots and once in dicots (Zhao, M. et al., 2014). Our analysis shows that another duplication event of this protein family occurred during the evolution of the Brassicaceae family, resulting in the evolution of the DA1- family clade II proteins which in *A. thaliana* contains DA1 family proteins DAR3-7. This result further sub-divides this family, adding a Brassicaceae specific clade. Furthermore, we identified a third clade that has evolved in the Rosids and contains the LIM-Peptidase containing TNLs identified from *Prunus persica* as well as proteins identified in *Theobroma cacao.* Therefore, the LIM-Peptidase protein family has undergone several duplication events in different plant lineages that have resulted in LIM-Peptidase containing resistance genes evolving on multiple occasions. Although, there aren't as of yet any LIM-Peptidase containing resistance genes identified from the monocot lineage.

The two LIM-peptidase domain containing resistance genes identified in *A. thaliana* both cause resistance to phytopathogen *A. candida* race 4 isolate AcEM2 and it is likely that other LIM-peptidase domain containing resistance genes identified in the Brassicaceae are also active against *A. candida* isolates. However, *A. candida* has a host range that is mainly confined to the Brassicaceae family (Saharan et al., 2014). Therefore, it is unlikely that LIM-peptidase containing resistance genes identified in non-Brassicaceae species are active against *A. candida* and instead confer resistance to another phytopathogen. This finding suggests that the LIM-Peptidase domain is a common target of multiple phytopathogens and that the ability of plants to recognise the perturbance that phytopathogens cause to the LIM-Peptidase domain is a mechanism that can confer resistance to multiple LIM-Peptidase targeting phytopathogens.

Brassicaceae clade I DA1 family proteins are important regulators of cell and organ size through their interaction and cleavage of E3 ligases and transcription factors (Li, Yunhai et al., 2008; Xia, T. et al., 2013; Du, Liang et al., 2014; Peng, Yuancheng et al., 2015; Vanhaeren et al., 2017). Clade II proteins, had no known function other than the involvement of WRR5B in autoimmunity and chilling sensitivity (Yang et al., 2010; Xu et al., 2015). Here we show that DA1 family clade II proteins WRR5B and WRR7 are key regulators of immunity to *A. candida* in *A. thaliana* and that DAR6 is important in regulating developmental programmes in *A. thaliana* and is particularly important in regulating leaf proliferation and the transition to flowering, particularly in the formation of the primary bolt. This function of DAR6 is novel to the Brassicaceae family and has evolved during the evolution of this family. We observed no phenotypic defects when overexpressing DAR3 and DAR7 and put forward the hypothesis that these two proteins are acting as decoys, recognising the presence of an as yet unidentified *A. candida* effector which is recognised by an associated resistance gene.

Although, it seems logical that the LIM-Peptidase domain of WRR5B and WRR7 interacts with an *A. candida* effector either directly or indirectly, no known effector has yet been identified. Without knowing what the exact *A. candida* effector is that interacts with the LIM-Peptidase domains of WRR5B and WRR7 then the full mechanistic action of these pathways can't be fully elucidated. Different *A. candida* races show differing abilities to infect *A. thaliana* accessions and some can infect the Col-0 ecotype that is resistant to AcEM2 (Borhan, M. Hossein et al., 2008; McMullan et al., 2015; Prince et al., 2017; Cevik et al., 2019). Therefore, for AcEM2 to be recognised by the WRR5 and WRR7 resistance mechanisms, this isolate must have gained a mutation in an effector, making it compatible with WRR5 and WRR7 recognition or lost a factor that can mask the effectors presence from these two resistance mechanisms. To fully understand the *A. candida-A. thaliana* pathosystem we need to identify the interaction that occurs between the pathogen and host. Until the effector from AcEM2 that interacts with WRR5B and WRR7 is identified, the exact role WRR5B and WRR7 and their integrated LIM-Peptidase domains play in immunity will remain uncertain.

White rust resistance genes in *Arabidopsis thaliana*

There were previously four known resistance genes in *A. thaliana* populations active against *A. candida* (Cevik et al., 2019). We have identified three additional resistance genes (*WRR5A, WRR5B* and WRR*7*) and two new mechanisms, mediated by *WRR5A* & *WRR5B* and *WRR7*, employed by *A. thaliana* to combat infection caused by *A. candida*. Therefore, *A. thaliana* populations harbour at least seven resistance genes active against one phytopathogenic species. The three resistance genes, that we identified are all present in the *A. thaliana* Col-0 genome and confer resistance to *A. candida* race 4 isolate AcEM2, as well as previously identified *A. candida* responsive resistance gene *WRR4A* (Borhan, M. Hossein et al., 2008). *WRR5A, WRR5B* and *WRR7* are involved in mediating *A. thaliana* resistance via two distinct mechanisms that are masked by the epistasis conferred by the highly sensitive mechanism mediated by *WRR4A*. Therefore, the presence of genomes harbouring stacked NLRs active against one pathogen with different epistatic redundancies offers an interesting insight into how NLRs can rapidly evolve in a population. If NLRs active against one phytopathogenic species are transmitted into a new population that already has an active resistance gene against the pathogen strain, then they will become functionally redundant and masked by epistasis. Therefore, the redundantly masked resistance gene will be released from its selective constraints and will mutate at increased rates, facilitating the production of new allelic variants of NLRs that can evolve novel functionality during host-pathogen co-evolution. In addition, *A. thaliana* populations maintain different combinations of resistance genes, and this variation is enough to cause resistance. Some of these populations maintain only single resistance genes e.g. Hi-0 (Cevik et al., 2019), whereas others maintain multiple resistance genes such as Col-0 (Fig 3.2). Others maintain multiple *WRR* genes capable of conferring resistance to some *A. candida* races but are susceptible to other physiological races of the same pathogen e.g. the Ws-2 genome encodes *WRR4B* and *WRR12* that confer resistance to *A. candida* races Ac2V and AcBoT (Cevik et al., 2019) but is susceptible to AcEM2 (Fig 3.3). This stratification of resistance genes across populations allows one species to maintain multiple resistance mechanisms without high evolutionary costs. As one population only has to maintain one or two mechanisms and the population as a whole can

therefore maintain many more. This dynamic community population structure allows ecological systems to maintain durable resistance against one pathogen. Therefore, a community genetics model may be one that we wish to employ in agricultural systems instead of generating one or two super varieties of crops that have multiple stacked traits. Super varieties containing multiple stacked resistance genes also have the drawback of generating extreme selective pressures upon pathogens that could lead to a super resistant pathogen strain that would undergo a selective sweep, risking high levels of crop losses. Therefore, it is likely to be more durable to generate multiple crop lines containing single or limited resistance gene stacks that are deployed in communities rather than generating crop varieties with many stacked NLRs.

Evolution of the LIM-Peptidase integrated decoy domain

Integrated domains are of increasing interest for crop breeding approaches (Tamborski and Krasileva, 2020). However, our understanding of how resistance genes evolve to encode integrated domains and subsequently evolve functionality is still not very well understood. NLR clades containing integrated domains are not evenly distributed in the genomes and some clades of NLRs such as the Major integration clade 1 (MIC1) from the Poaceae family are highly enriched for integrated domains (Bailey et al., 2018). Therefore, some of these NLR clades 'shuffle' integrated domains into their architectures, generating new sensing specificities (Tamborski and Krasileva, 2020). In this model, the integrated domains translocate into the NLR, however it is also plausible that NLRs themselves are the motile element in the genome and fuse to novel integrated domains by NLR translocation. The exact evolutionary mechanisms that are involved in NLRs gaining novel domain fusions is still largely unknown. However, we understand even less about the process of how a novel integrated domain fusion gains recognition specificity with their respective pathogens.

The LIM-peptidase domain fusion is found in Brassicaceae genomes in homologs of the *A. thaliana WRR5B* and *WRR7* genes that we have shown to cause resistance against the biotrophic Oomycete pathogen *A. candida*. In addition, this integrated domain is encoded by resistance genes from the Rosaceae, Fabaceae and Vitaceae families (Srivastava and Verma, 2015; Sarris, P. F. et al., 2016). We were able to

show, that this integration event occurred at least four times and occurred twice during the evolution of the Brassicaceae (Fig 6.12). Furthermore, we were able to track homologs of *WRR5B* and *WRR7* in closely related species *Eutrema salsuginium* and *Carica papaya* that did not contain integrated LIM-Peptidase and may represent genes that resemble the ancestral form of *WRR5B* and *WRR7.* Intriguingly, the *WRR5B* homolog identified in *E. salsuginium* was located in a syntenic loci with the *A. thaliana WRR5B* gene and therefore represents an evolutionary scenario where the integrated domain translocated into the locus containing the NLR. In contrast, the *WRR7* CC_R homolog identified in *C. papaya* was located in a syntenic region to the *A. thaliana NRG* genes and not the locus containing *WRR7*. Therefore, this represents an evolutionary event where the NLR itself translocated into the locus that originally contained the integrated domain. Therefore, we provide evidence that the integrated domain or the NLR can translocate, giving rise to functional NLRs containing integrated domains that can recognise a plant pathogen. How these NLR fusions evolved to confer resistance to *A. candida* is still unknown. However, studying the evolution of these two integration events and revealing how they gained recognition specificity to *A. candida* race 4 isolate AcEM2 could shed light on the co-evolutionary processes that give rise to functional NLR-ID fusions.

Deploying the integrated domain in crop varieties

In order to engineer crop varieties with durable NLR based defence mechanisms against phytopathogens, we need to be able to engineer and innovate novel resistance genes. To date, the majority of successful engineering approaches taken to developing crop varieties with new arsenals of NLRs has involved moving existing resistance from one species or cultivar into another, although the majority of successful transfers have involved the transfer of cell surface receptors rather than NLRs (Rodriguez-Moreno, Song and Thomma, 2017). More recently approaches to alter the binding structure of integrated domains, such as the HMA domain of *O. sativa* NLR PikP have had success in increasing the binding affinity of NLRs to different effector variants (De la Concepcion et al., 2019). However, no engineering approach utilising a novel NLR construct has been successful in field trials. It is becoming increasingly clear that plants utilise a variety of NLR based mechanisms to

stimulate ETI. These include using NLRs that utilise different signalling mechanisms, NLRs that operate as heterodimers and NLRs with integrated domains. In addition, one species can have multiple populations that retain different NLR combinations that confer resistance to a single phytopathogen. Therefore, we are only just beginning to appreciate the complexity of NLR based immunity and how plants evolve new NLRs or NLR derivatives to combat biotrophic phytopathogens.

One of the most interesting avenues of research for NLR based engineering approaches is to exploit integrated decoy domains to engineer novel recognition specificities to pathogen effectors. So far integrated decoy domains have been engineered to have increased affinity for a wider range of effectors but the generation of a new NLR with a novel integrated decoy domain that can recognise a pathogen effector has not been achieved (De la Concepcion et al., 2019; Tamborski and Krasileva, 2020).

In order to engineer a resistance gene with a novel integrated domain you need to have a receptive NLR 'base' architecture that can receive an integrated domain without perturbing its immune signalling function. However, most integrated domain NLRs are being found to require a second helper NLR to execute immune function, therefore you would need to transfer both sensor and helper NLRs into the engineered host to activate resistance (Césari et al., 2014; Rodriguez-Moreno, Song and Thomma, 2017). *WRR7* is an unusual CC_R type resistance gene, in that it contains an integrated LIM-Peptidase domain but seemingly has no associated helper NLR. This unusual CC_R-LIM-peptidase structure was also observed to have independently evolved in *Cicer arietinum* and a similar TN-LIM-peptidase fusion was observed in Rosaceae species *Malus domestica* and *Prunus persica*. Therefore, as a CC_R protein that is highly similar to NRG helper NLRs, WRR7 may have evolved the ability to both sense the presence of *A. candida* and execute immune signalling independently of other NLRs. This makes WRR7 an interesting candidate 'base' NLR architecture for integrated decoy engineering approaches. However, we have shown through domain swapping experiments using WRR7 as the base architecture that even subbing highly similar domains into NLR architectures cannot elicit similar immune responses to the native NLR. Although, we were able to elicit mild autoimmune responses with domain swaps of the WRR7 LIM-Peptidase domains for the LIM-peptidase domains of WRR5B and DAR3. Therefore, for a

novel NLR fusion with an integrated domain to become functional the NLR and integrated domain have to undergo specific host-pathogen co-evolution for the domain to become sensitive enough to pathogen effectors to elicit an immune response.

Therefore, engineering approaches utilising the integrated decoy model of NLR activation are not straight forward and a scenario where one NLR architecture can be used to sub in multiple diverse decoy domains seems improbable. However, our current understanding of how NLRs evolve to contain integrated decoy domains is still poor and we have shown that integrated decoy domains can be gained through the translocation of either the NLR or the integrated domain (Fig 6.13 and 6.15). Our understanding of how these NLR fusions subsequently acquire functionality is lacking. Until we fully understand how functionality arises following a domain integration, the generation of NLRs with novel functional integrated domains will be hard to produce and we will have to rely on mining natural populations for functional NLRs to use in crops.

Conclusion

The *A. thaliana* Col-0 genome harbours four resistance genes and three resistance mechanisms that confer resistance to *A. candida* race 4 isolate AcEM2. The WRR5A and WRR5B mechanism operates by the sensor-helper model of NLR activation with WRR5B acting as the sensor and WRR5A acting as the executioner, although unusually the P-loop region of both of these proteins is required for immune activation. The *WRR7* mediated immune signalling response to *A. candida* seemingly acts independently of other NLRs and WRR7 may have evolved the ability to sense the presence of its associated phytopathogen and execute immune signalling by itself. All three NLRs studied (WRR5A, WRR5B and WRR7) have shown plasma membrane localisation and WRR5A and WRR5B are also localised to the nucleus and WRR5A is enriched in the nucleus of *A. thaliana* proceeding infection by *A. candida*. We have identified several novel genes that are involved in the *WRR7* mediated immune pathway (*CHR4, MAC7, MAP3Kδ4* and *CAMTA2*) as well as identifying several proteins that interact with the WRR5 complex including LOX2, PIP2 and BRM. In addition, we were able to show that WRR5B and WRR7 gained their integrated LIM-Peptidase domains during the evolution of the Brassicaceae family via two different integration events. We then attempted to

reproduce the evolution of WRR7 using domain swapping experiments to introduce different LIM-Peptidase domains from the DA1 family onto the base architecture of WRR7. We were unable to recreate resistance to *A. candida* with these fusion proteins. Therefore, although in theory generation of novel NLRs containing integrated domains is an enticing prospect, in reality the specificity of NLR interactions with effector proteins makes this goal incredibly challenging to pull off.

Bibliography

Aalfs, J.D. and Kingston, R.E., 2000. What does 'chromatin remodeling' mean? *Trends in Biochemical Sciences,* 25(11), pp. 548-555.

Albrecht, T. and Argueso, C.T., 2017. Should I fight or should I grow now? The role of cytokinins in plant growth and immunity and in the growth–defence trade-off. *Annals of Botany,* 119(5), pp. 725-735.

Alexandratos, N. and Bruinsma, J., 2012. *World agriculture towards 2030/2050: the 2012 revision.*

Anders, S., Pyl, P.T. and Huber, W., 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England),* 31(2), pp. 166-169.

Arae, T., Morita, K., Imahori, R., Suzuki, Y., Yasuda, S., Sato, T., Yamaguchi, J. and Chiba, Y., 2019. Identification of Arabidopsis CCR4-NOT Complexes with Pumilio RNA-Binding Proteins, APUM5 and APUM2. *Plant and Cell Physiology,* 60(9), pp. 2015-2025.

Ashikawa, I., Hayashi, N., Yamane, H., Kanamori, H., Wu, J., Matsumoto, T., Ono, K. and Yano, M., 2008. Two Adjacent Nucleotide-Binding Site–Leucine-Rich Repeat Class Genes Are Required to Confer Pikm-Specific Rice Blast Resistance. *Genetics,* 180(4), pp. 2267-2276.

Baggs, E., Dagdas, G. and Krasileva, K.V., 2017. NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Current Opinion in Plant Biology,* 38, pp. 59-67.

Bailey, P.C., Schudoma, C., Jackson, W., Baggs, E., Dagdas, G., Haerty, W., Moscou, M. and Krasileva, K.V., 2017. Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. *bioRxiv*.

Bailey, P.C., Schudoma, C., Jackson, W., Baggs, E., Dagdas, G., Haerty, W., Moscou, M. and Krasileva, K.V., 2018. Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. *Genome Biology,* 19(1), p. 23.

Behrens, F.H., Schenke, D., Hossain, R., Ye, W., Schemmel, M., Bergmann, T., Häder, C., Zhao, Y., Ladewig, L., Zhu, W. and Cai, D., 2019. Suppression of abscisic acid biosynthesis at the early infection stage of Verticillium longisporum in oilseed rape (Brassica napus). *Molecular Plant Pathology,* 20(12), pp. 1645-1661.

Bent, A., 2006. Arabidopsis thaliana Floral Dip Transformation Method. In: K. Wang, ed. *Agrobacterium Protocols.* Totowa, NJ: Humana Press, pp. 87-104.

Bentham, A., Burdett, H., Anderson, P.A., Williams, S.J. and Kobe, B., 2016. Animal NLRs provide structural insights into plant NLR function. *Ann Bot,* p. mcw171.

Bentham, A.R., Zdrzałek, R., De la Concepcion, J.C. and Banfield, M.J., 2018. Uncoiling CNLs: Structure/function approaches to understanding CC domain function in plant NLRs. *Plant and Cell Physiology,* pp. pcy185-pcy185.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. and Huala, E., 2015. The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. *genesis,* 53(8), pp. 474-485.

Berens, M.L., Berry, H.M., Mine, A., Argueso, C.T. and Tsuda, K., 2017. Evolution of Hormone Signaling Networks in Plant Defense. *Annual Review of Phytopathology,* 55(1), pp. 401-425.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L. and Lander, E.S., 2006. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell,* 125(2), pp. 315-326.

Bi, D., Johnson, K.C.M., Zhu, Z., Huang, Y., Chen, F., Zhang, Y. and Li, X., 2011. Mutations in an atypical TIR-NB-LRR-LIM resistance protein confer autoimmunity. *Frontiers in Plant Science,* 2.

Bi, G. and Zhou, J.-M., 2017. MAP Kinase Signaling Pathways: A Hub of Plant-Microbe Interactions. *Cell Host & Microbe,* 21(3), pp. 270-273.

Bialas, A., Zess, E.K., De la Concepcion, J.C., Franceschetti, M., Pennington, H.G., Yoshida, K., Upson, J.L., Chanclud, E., Wu, C.-H., Langner, T., Maqbool, A., Varden, F.A., Derevnina, L., Belhaj, K., Fujisaki, K., Saitoh, H., Terauchi, R., Banfield, M.J. and Kamoun, S., 2017. Lessons in effector and NLR biology of plant-microbe systems. *bioRxiv*.

Bonardi, V., Tang, S., Stallmann, A., Roberts, M., Cherkis, K. and Dangl, J.L., 2011. Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. *Proceedings of the National Academy of Sciences,* 108(39), pp. 16463-16468.

Borhan, M.H., Gulden, S., Rimmer, S.R., Gunn, N., Cooper, A., Tör, M. and Holub, E.B., 2008. WRR4 encodes a TIR- NB- LRR protein that confers broad- spectrum white rust resistance in Arabidopsis thaliana to four physiological races of Albugo candida. *Molecular Plant-Microbe Interactions,* 21(6), pp. 757-768.

Borhan, M.H., Holub, E.B., Kindrachuk, C., Omidi, M., Bozorgmanesh-Frad, G. and Rimmer, S.R., 2010. WRR4, a broad-spectrum TIR-NB-LRR gene from Arabidopsis thaliana that confers white rust resistance in transgenic oilseed brassica crops. *Molecular Plant Pathology,* 11(2), pp. 283-291.

Borrelli, G.M., Mazzucotelli, E., Marone, D., Crosatti, C., Michelotti, V., Valè, G. and Mastrangelo, A.M., 2018. Regulation and Evolution of NLR Genes: A Close Interconnection for Plant Immunity. *International journal of molecular sciences,* 19(6), p. 1662.

Bouché, N., Scharlat, A., Snedden, W., Bouchez, D. and Fromm, H., 2002. A Novel Family of Calmodulin-binding Transcription Activators in Multicellular Organisms. *Journal of Biological Chemistry,* 277(24), pp. 21851-21861.

Byrt, C.S., Zhao, M., Kourghi, M., Bose, J., Henderson, S.W., Qiu, J., Gilliham, M., Schultz, C., Schwarz, M., Ramesh, S.A., Yool, A. and Tyerman, S., 2017. Non-selective cation channel activity of aquaporin AtPIP2;1 regulated by Ca2+ and pH. *Plant, Cell & Environment,* 40(6), pp. 802-815.

Bökel, C., 2008. EMS Screens. In: C. Dahmann, ed. *Drosophila: Methods and Protocols.* Totowa, NJ: Humana Press, pp. 119-138.

Büttner, D. and He, S.Y., 2009. Type III Protein Secretion in Plant Pathogenic Bacteria. *Plant Physiology,* 150(4), pp. 1656-1664.

Camejo, D., Guzmán-Cedeño, Á. and Moreno, A., 2016. Reactive oxygen species, essential molecules, during plant–pathogen interactions. *Plant Physiology and Biochemistry,* 103, pp. 10-23.

Canto, T., Uhrig, J.F., Swanson, M., Wright, K.M. and MacFarlane, S.A., 2006. Translocation of Tomato bushy stunt virus P19 protein into the nucleus by ALY proteins compromises its silencing suppressor activity. *Journal of virology,* 80(18), pp. 9064-9072.

Carotenuto, G., Chabaud, M., Miyata, K., Capozzi, M., Takeda, N., Kaku, H., Shibuya, N., Nakagawa, T., Barker, D.G. and Genre, A., 2017. The rice LysM receptor-like kinase OsCERK1 is required for the perception of short-chain chitin oligomers in arbuscular mycorrhizal signaling. *New Phytologist,* 214(4), pp. 1440-1446.

Carter, B., Bishop, B., Ho, K.K., Huang, R., Jia, W., Zhang, H., Pascuzzi, P.E., Deal, R.B. and Ogas, J., 2018. The Chromatin Remodelers PKL and PIE1 Act in an Epigenetic Pathway That Determines H3K27me3 Homeostasis in Arabidopsis. *The Plant Cell,* 30(6), pp. 1337-1352.

Castel, B., Ngou, P.-M., Cevik, V., Redkar, A., Kim, D.-S., Yang, Y., Ding, P. and Jones, J.D.G., 2018. Diverse NLR immune receptors activate defence via the RPW8-NLR NRG1. *New Phytologist,* 0(ja).

Cesari, S., 2017. Multiple strategies for pathogen perception by plant immune receptors. *New Phytologist,* 219(1), pp. 17-24.

Cesari, S., Kroj, T., Bernoux, M., Dodds, P.N. and Moncuquet, P., 2014. A novel conserved mechanism for plant NLR protein pairs: The "integrated decoy" hypothesis. *Frontiers in Plant Science,* 5.

Cesari, S., Thilliez, G., Ribot, C., Chalvon, V., Michel, C., Jauneau, A., Rivas, S., Alaux, L., Kanzaki, H., Okuyama, Y., Morel, J.-B., Fournier, E., Tharreau, D., Terauchi, R. and Kroj, T., 2013. The Rice Resistance Protein Pair RGA4/RGA5 Recognizes the Magnaporthe oryzae Effectors AVR-Pia and AVR1-CO39 by Direct Binding. *The Plant Cell,* 25(4), pp. 1463-1481.

Cevik, V., Boutrot, F., Apel, W., Robert-Seilaniantz, A., Furzer, O.J., Redkar, A., Castel, B., Kover, P.X., Prince, D.C., Holub, E.B. and Jones, J.D.G., 2019. Transgressive segregation reveals mechanisms of *Arabidopsis* immunity to *Brassica*-infecting races of white rust (*Albugo candida*). *Proceedings of the National Academy of Sciences,* 116(7), pp. 2767-2773.

Charpentier, M., Sun, J., Martins, T.V., Radhakrishnan, G.V., Findlay, K., Soumpourou, E., Thouin, J., Véry, A.-A., Sanders, D., Morris, R.J. and Oldroyd, G.E.D., 2016. Nuclear-localized cyclic nucleotide–gated channels mediate symbiotic calcium oscillations. *Science,* 352(6289), pp. 1102-1105.

Chen, K., Li, G.-J., Bressan, R.A., Song, C.-P., Zhu, J.-K. and Zhao, Y., 2020. Abscisic acid dynamics, signaling, and functions in plants. *Journal of Integrative Plant Biology,* 62(1), pp. 25-54.

Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M.L., Qu, X.-Q., Guo, W.-J., Kim, J.-G., Underwood, W., Chaudhuri, B., Chermak, D., Antony, G., White, F.F., Somerville, S.C., Mudgett, M.B. and Frommer, W.B., 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature,* 468(7323), pp. 527-532.

Chen, X., Liu, X., Zhao, Y. and Zhou, D.-X., 2015. Histone H3K4me3 and H3K27me3 regulatory genes control stable transmission of an epimutation in rice. *Scientific Reports,* 5(1), p. 13251.

Chen, Z., Zheng, Z., Huang, J., Lai, Z. and Fan, B., 2009. Biosynthesis of salicylic acid in plants. *Plant Signalling & Behaviour,* 4(6).

Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S., Liu, B., Li, P., Hua, W. and Wang, X., 2011. BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biology,* 11(1), p. 136.

Cheval, C., Aldon, D., Galaud, J.-P. and Ranty, B., 2013. Calcium/calmodulin-mediated regulation of plant immunity. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1833(7), pp. 1766-1771.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly,* 6(2), pp. 80-92.

Cooper, A.J., Latunde-Dada, A.O., Woods-Tör, A., Lynn, J., Lucas, J.A., Crute, I.R. and Holub, E.B., 2008. Basic Compatibility of Albugo candida in Arabidopsis thaliana and Brassica juncea Causes Broad-Spectrum Suppression of Innate Immunity. *Molecular Plant-Microbe Interactions,* 21(6), pp. 745-756.

Couto, D. and Zipfel, C., 2016. Regulation of pattern recognition receptor signalling in plants. *Nature Reviews Immunology,* 16(9), pp. 537-552.

Cui, H., Gobbato, E., Kracher, B., Qiu, J., Bautor, J. and Parker, J.E., 2017. A core function of EDS1 with PAD4 is to protect the salicylic acid defense sector in Arabidopsis immunity. *New Phytologist,* 213(4), pp. 1802-1817.

Cui, H., Tsuda, K. and Parker, J.E., 2015. Effector- Triggered Immunity: From Pathogen Perception to Robust Defense. *Annu. Rev. Plant Biol.*, pp. 487-511.

Czech, M.P., 2000. PIP2 and PIP3: Complex Roles at the Cell Surface. *Cell,* 100(6), pp. 603-606.

Césari, S., Kanzaki, H., Fujiwara, T., Bernoux, M., Chalvon, V., Kawano, Y., Shimamoto, K., Dodds, P., Terauchi, R. and Kroj, T., 2014. The NB - LRR proteins RGA 4 and RGA 5 interact functionally and physically to confer disease resistance. *EMBO Journal,* 33(17), pp. 1941-1959.

Das, G. and Rao, G.J.N., 2015. Molecular marker assisted gene stacking for biotic and abiotic stress resistance genes in an elite rice cultivar. *Frontiers in Plant Science,* 6(698).

Day, B., Dahlbeck, D. and Staskawicz, B.J., 2006a. NDR1 Interaction with RIN4 Mediates the Differential Activation of Multiple Disease Resistance Pathways in Arabidopsis. *The Plant Cell,* 18(10), pp. 2782-2791.

Day, B., Dahlbeck, D. and Staskawicz, B.J., 2006b. NDR1 Interaction with RIN4 Mediates the Differential Activation of Multiple Disease Resistance Pathways in *Arabidopsis*. *The Plant Cell*, 18(10), pp. 2782-2791.

De, I., Bessonov, S., Hofele, R., dos Santos, K., Will, C.L., Urlaub, H., Lührmann, R. and Pena, V., 2015. The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. *Nature Structural & Molecular Biology*, 22(2), pp. 138-144.

De la Concepcion, J.C., Franceschetti, M., MacLean, D., Terauchi, R., Kamoun, S. and Banfield, M.J., 2019. Protein engineering expands the effector recognition profile of a rice NLR immune receptor. *eLife*, 8, p. e47713.

de Torres-Zabala, M., Truman, W., Bennett, M.H., Lafforgue, G., Mansfield, J.W., Rodriguez Egea, P., Bögre, L. and Grant, M., 2007. Pseudomonas syringae pv. tomato hijacks the Arabidopsis abscisic acid signalling pathway to cause disease. *The EMBO Journal*, 26(5), pp. 1434-1443.

del Pozo, O., Pedley, K.F. and Martin, G.B., 2004. MAPKKKalpha is a positive regulator of cell death associated with both plant immunity and disease. *Embo j*, 23(15), pp. 3072-3082.

Derevnina, L., Dagdas, Y.F., De la Concepcion, J.C., Bialas, A., Kellner, R., Petre, B., Domazakis, E., Du, J., Wu, C.-H., Lin, X., Aguilera-Galvez, C., Cruz-Mireles, N., Vleeshouwers, V.G.A.A. and Kamoun, S., 2016. Nine things to know about elicitins. *New Phytologist*, 212(4), pp. 888-895.

Deslandes, L., Olivier, J., Peeters, N., Feng, D.X., Khounlotham, M., Boucher, C., Somssich, I., Genin, S. and Marco, Y., 2003. Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proceedings of the National Academy of Sciences*, 100(13), pp. 8024-8029.

Ding, P. and Redkar, A., 2018. Pathogens Suppress Host Transcription Factors for Rampant Proliferation. *Trends in Plant Science*, 23(11), pp. 950-953.

Dodd, A.N., Kudla, J. and Sanders, D., 2010. The Language of Calcium Signaling. *Annual Review of Plant Biology*, 61(1), pp. 593-620.

Dodds, P.N. and Rathjen, J.P., 2010. Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature Reviews Genetics*, 11, p. 539.

Dong, H., Dumenil, J., Lu, F.-H., Na, L., Vanhaeren, H., Naumann, C., Klecker, M., Prior, R., Smith, C., McKenzie, N., Saalbach, G., Chen, L., Xia, T., Gonzalez, N., Seguela, M., Inzé, D., Dissmeyer, N., Li, Y. and Bevan, M.W., 2017. Ubiquitylation activates a peptidase that promotes cleavage and destabilization of its activating E3 ligases and diverse growth regulatory proteins to limit cell proliferation in Arabidopsis. *Genes & Development*, 31(2), pp. 197-208.

Dong, J., Chen, C. and Chen, Z., 2003. Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. *Plant Molecular Biology*, 51(1), pp. 21-37.

Dong, O.X., Tong, M., Bonardi, V., El Kasmi, F., Woloshen, V., Wünsch, L.K., Dangl, J.L. and Li, X., 2016. TNL-mediated immunity in Arabidopsis requires complex regulation of the redundant ADR1 gene family. *New Phytologist*, 210(3), pp. 960-973.

Du, L., Ali, G.S., Simons, K.A., Hou, J., Yang, T., Reddy, A.S.N. and Poovaiah, B.W., 2009. Ca2+/calmodulin regulates salicylic-acid-mediated plant immunity. *Nature*, 457, p. 1154.

Du, L., Li, N., Chen, L., Xu, Y., Li, Y., Zhang, Y., Li, C. and Li, Y., 2014. The ubiquitin receptor DA1 regulates seed and organ size by modulating the stability of the ubiquitin-specific protease UBP15/SOD2 in Arabidopsis. *The Plant cell*, 26(2), p. 665.

Duszyn, M., Świeżawska, B., Szmidt-Jaworska, A. and Jaworski, K., 2019. Cyclic nucleotide gated channels (CNGCs) in plant signalling—Current knowledge and perspectives. *Journal of Plant Physiology*, 241, p. 153035.

Dyrka, W., Lamacchia, M., Durrens, P., Kobe, B., Daskalov, A., Paoletti, M., Sherman, D.J. and Saupe, S.J., 2014. Diversity and Variability of NOD-Like Receptors in Fungi. *Genome Biology and Evolution*, 6(12), pp. 3137-3158.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), p. 1792.

Eitas, T.K. and Dangl, J.L., 2010. NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Current Opinion in Plant Biology*, 13(4), pp. 472-477.

Eliasson, Å., Gass, N., Mundel, C., Baltz, R., Kräuter, R., Evrard, J.L. and Steinmetz, A., 2000. Molecular and expression analysis of a LIM protein gene family from flowering plants. *Molecular and General Genetics MGG*, 264(3), pp. 257-267.

Fernández-Bautista, N., Domínguez-Núñez, J.A., Moreno, M.M.C. and Berrocal-Lobo, M., 2016. Plant Tissue Trypan Blue Staining During Phytopathogen Infection. *Bio-protocol*, 6(24), p. e2078.

Finkelstein, R., 2013. Abscisic Acid synthesis and response. *The arabidopsis book*, 11, pp. e0166-e0166.

Finkler, A., Ashery-Padan, R. and Fromm, H., 2007. CAMTAs: Calmodulin-binding transcription activators from plants to human. *FEBS Letters*, 581(21), pp. 3893-3898.

Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., Bateman, A. and Eddy, S.R., 2015. HMMER web server: 2015 update. *Nucleic Acids Research*, 43(W1), pp. W30-W38.

Fischer, C., Kugler, A., Hoth, S. and Dietrich, P., 2013. An IQ domain mediates the interaction with calmodulin in a plant cyclic nucleotide-gated channel. *Plant & cell physiology*, 54(4), pp. 573-584.

Flor, H.H., 1971. Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology*, 9(1), pp. 275-296.

Fujisaki, K., Abe, Y., Kanzaki, E., Ito, K., Utsushi, H., Saitoh, H., Białas, A., Banfield, M., Kamoun, S. and Terauchi, R., 2017. An unconventional NOI/RIN4 domain of a rice NLR protein binds host EXO70 protein to confer fungal immunity. *bioRxiv*.

Galon, Y., Snir, O. and Fromm, H., 2010. How calmodulin binding transcription activators (CAMTAs) mediate auxin responses. *Plant signaling & behavior*, 5(10), pp. 1311-1314.

Gantner, J., Ordon, J., Kretschmer, C., Guerois, R. and Stuttmann, J., 2019. An EDS1-SAG101 Complex Is Essential for TNL-Mediated Immunity in *Nicotiana benthamiana*. *The Plant Cell*, 31(10), pp. 2456-2474.

Gao, Q.-M., Zhu, S., Kachroo, P. and Kachroo, A., 2015. Signal regulators of systemic acquired resistance. *Frontiers in Plant Science*, 6(228).

Gao, X., Chen, X., Lin, W., Chen, S., Lu, D., Niu, Y., Li, L., Cheng, C., McCormack, M., Sheen, J., Shan, L. and He, P., 2013. Bifurcation of Arabidopsis NLR Immune Signaling via Ca2+-Dependent Protein Kinases. *PLOS Pathogens*, 9(1), p. e1003127.

Gao, Y., Wang, W., Zhang, T., Gong, Z., Zhao, H. and Han, G.-Z., 2018. Out of Water: The Origin and Early Diversification of Plant R-Genes. *Plant Physiology*.

Ge, S.X., Son, E.W. and Yao, R., 2018. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics*, 19(1), p. 534.

Genre, A., Chabaud, M., Balzergue, C., Puech-Pagès, V., Novero, M., Rey, T., Fournier, J., Rochange, S., Bécard, G., Bonfante, P. and Barker, D.G., 2013. Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca2+ spiking in Medicago truncatula roots and their production is enhanced by strigolactone. *New Phytologist*, 198(1), pp. 190-202.

Gentry, M. and Hennig, L., 2014. Remodelling chromatin to shape development of plants. *Experimental Cell Research*, 321(1), pp. 40-46.

Gibbs, H.K., Ruesch, A.S., Achard, F., Clayton, M.K., Holmgren, P., Ramankutty, N. and Foley, J.A., 2010. Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. *Proceedings of the National Academy of Sciences*, 107(38), pp. 16732-16737.

Glazebrook, J., 2005. Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annual Review of Phytopathology*, 43(1), pp. 205-227.

González-Fuente, M., Carrère, S., Monachello, D., Marsella, B.G., Cazalé, A.-C., Zischek, C., Mitra, R.M., Rezé, N., Cottret, L., Mukhtar, M.S., Lurin, C., Noel, L.D. and Peeters, N., 2019. EffectorK, a comprehensive resource to mine for pathogen effector targets in the Arabidopsis proteome. *bioRxiv*, p. 2019.2012.2016.878074.

Gouy, M., Guindon, S. and Gascuel, O., 2009. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution*, 27(2), pp. 221-224.

Grant, M., Brown, I., Adams, S., Knight, M., Ainslie, A. and Mansfield, J., 2000. The RPM1 plant disease resistance gene facilitates a rapid and sustained increase in cytosolic calcium that is necessary for the oxidative burst and hypersensitive cell death. *The Plant Journal*, 23(4), pp. 441-450.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O., 2010. New Algorithms and Methods to Estimate Maximum- Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, 59(3), pp. 307-321.

Gómez-Gómez, L. and Boller, T., 2000. FLS2: An LRR Receptor–like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis. *Molecular Cell*, 5(6), pp. 1003-1011.

Göhre, V., Spallek, T., Häweker, H., Mersmann, S., Mentzel, T., Boller, T., de Torres, M., Mansfield, J.W. and Robatzek, S., 2008. Plant Pattern-Recognition Receptor FLS2 Is Directed for Degradation by the Bacterial Ubiquitin Ligase AvrPtoB. *Current Biology,* 18(23), pp. 1824-1832.

Ha, M., 2013. Understanding the chromatin remodeling code. *Plant Science,* 211, pp. 137-145.

Hander, T., Fernández-Fernández, Á.D., Kumpf, R.P., Willems, P., Schatowitz, H., Rombaut, D., Staes, A., Nolf, J., Pottie, R., Yao, P., Gonçalves, A., Pavie, B., Boller, T., Gevaert, K., Van Breusegem, F., Bartels, S. and Stael, S., 2019. Damage on plants activates Ca^{2+}-dependent metacaspases for release of immunomodulatory peptides. *Science,* 363(6433), p. eaar7486.

Haug-Baltzell, A., Stephens, S.A., Davey, S., Scheidegger, C.E. and Lyons, E., 2017. SynMap2 and SynMap3D: web-based whole-genome synteny browsers. *Bioinformatics,* 33(14), pp. 2197-2198.

Hennig, L. and Derkacheva, M., 2009. Diversity of Polycomb group complexes in plants: same rules, different players? *Trends in Genetics,* 25(9), pp. 414-423.

Herzel, L., Ottoz, D.S.M., Alpert, T. and Neugebauer, K.M., 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nature Reviews Molecular Cell Biology,* 18(10), pp. 637-650.

Holt, B.F., Belkhadir, Y. and Dangl, J.L., 2005. Antagonistic Control of Disease Resistance Protein Stability in the Plant Immune System. *Science,* 309(5736), pp. 929-932.

Holub, E.B., Brose, E., Tor, M., Clay, C., Crute, I.R. and Beynon, J.L., 1995. Phenotypic and genotypic variation in the interaction between Arabidopsis thaliana and Albugo candida. *Mol Plant Microbe Interact,* 8(6), pp. 916-928.

Hong, S.M., Bahn, S.C., Lyu, A., Jung, H.S. and Ahn, J.H., 2010. Identification and Testing of Superior Reference Genes for a Starting Pool of Transcript Normalization in Arabidopsis. *Plant and Cell Physiology,* 51(10), pp. 1694-1706.

Horsefield, S., Burdett, H., Zhang, X., Manik, M.K., Shi, Y., Chen, J., Qi, T., Gilley, J., Lai, J.-S., Rank, M.X., Casey, L.W., Gu, W., Ericsson, D.J., Foley, G., Hughes, R.O., Bosanac, T., von Itzstein, M., Rathjen, J.P., Nanson, J.D., Boden, M., Dry, I.B., Williams, S.J., Staskawicz, B.J., Coleman, M.P., Ve, T., Dodds, P.N. and Kobe, B., 2019. NAD^+ cleavage activity by animal and plant TIR domains in cell death pathways. *Science,* 365(6455), pp. 793-799.

Howe, E.A., Sinha, R., Schlauch, D. and Quackenbush, J., 2011. RNA-Seq analysis in MeV. *Bioinformatics (Oxford, England),* 27(22), pp. 3209-3210.

Hu, G., Cui, K., Northrup, D., Liu, C., Wang, C., Tang, Q., Ge, K., Levens, D., Crane-Robinson, C. and Zhao, K., 2013. H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. *Cell stem cell,* 12(2), pp. 180-192.

Hu, Y., Lai, Y. and Zhu, D., 2014. Transcription regulation by CHD proteins to control plant development. *Frontiers in plant science,* 5, pp. 223-223.

Hu, Y., Liu, D., Zhong, X., Zhang, C., Zhang, Q. and Zhou, D.-X., 2012. CHD3 protein recognizes and regulates methylated histone H3 lysines 4 and 27 over a subset of targets in the rice genome. *Proceedings of the National Academy of Sciences,* 109(15), pp. 5773-5778.

Hu, Z., Yan, C., Liu, P., Huang, Z., Ma, R., Zhang, C., Wang, R., Zhang, Y., Martinon, F., Miao, D., Deng, H., Wang, J., Chang, J. and Chai, J., 2013. Crystal Structure of NLRC4 Reveals Its Autoinhibition Mechanism. *Science,* 341(6142), pp. 172-175.

Huang, S., Balgi, A., Pan, Y., Li, M., Zhang, X., Du, L., Zhou, M., Roberge, M. and Li, X., 2016. Identification of Methylosome Components as Negative Regulators of Plant Immunity Using Chemical Genetics. *Molecular Plant,* 9(12), pp. 1620-1633.

Hubert, D.A., Tornero, P., Belkhadir, Y., Krishna, P., Takahashi, A., Shirasu, K. and Dangl, J.L., 2003. Cytosolic HSP90 associates with and modulates the Arabidopsis RPM1 disease resistance protein. *The EMBO journal,* 22(21), pp. 5679-5689.

Huh, S.U., Cevik, V., Ding, P., Duxbury, Z., Ma, Y., Tomlinson, L., Sarris, P.F. and Jones, J.D.G., 2017. Protein-protein interactions in the RPS4/RRS1 immune receptor complex. *PLOS Pathogens,* 13(5), p. e1006376.

Hunt, L., Lerner, F. and Ziegler, M., 2004. NAD – new roles in signalling and gene regulation in plants. *New Phytologist,* 163(1), pp. 31-44.

Hyun, K., Jeon, J., Park, K. and Kim, J., 2017. Writing, erasing and reading histone lysine methylations. *Experimental & Molecular Medicine,* 49(4), pp. e324-e324.

Irieda, H., Inoue, Y., Mori, M., Yamada, K., Oshikawa, Y., Saitoh, H., Uemura, A., Terauchi, R., Kitakura, S., Kosaka, A., Singkaravanit-Ogawa, S. and Takano, Y., 2019. Conserved fungal effector suppresses PAMP-triggered immunity by targeting plant immune kinases. *Proceedings of the National Academy of Sciences,* 116(2), pp. 496-505.

Jacob, F., Kracher, B., Mine, A., Seyfferth, C., Blanvillain-Baufumé, S., Parker, J.E., Tsuda, K., Schulze-Lefert, P. and Maekawa, T., 2017. A dominant-interfering camta3 mutation compromises primary transcriptional outputs mediated by both cell surface and intracellular immune receptors in Arabidopsis thaliana. *New Phytologist,* pp. n/a-n/a.

Jambhekar, A., Dhall, A. and Shi, Y., 2019. Roles and regulation of histone methylation in animal development. *Nature Reviews Molecular Cell Biology,* 20(10), pp. 625-641.

Jia, T., Zhang, B., You, C., Zhang, Y., Zeng, L., Li, S., Johnson, K.C.M., Yu, B., Li, X. and Chen, X., 2017. The Arabidopsis MOS4-Associated Complex Promotes MicroRNA Biogenesis and Precursor Messenger RNA Splicing. *The Plant Cell,* 29(10), pp. 2626-2643.

Jiang, C.-J., Shimono, M., Sugano, S., Kojima, M., Yazawa, K., Yoshida, R., Inoue, H., Hayashi, N., Sakakibara, H. and Takatsuji, H., 2010. Abscisic Acid Interacts Antagonistically with Salicylic Acid Signaling Pathway in Rice–Magnaporthe grisea Interaction. *Molecular Plant-Microbe Interactions®,* 23(6), pp. 791-798.

Jing, Y., Lin, R. and Guo, Q., 2019. The Chromatin-Remodeling Factor PICKLE Antagonizes Polycomb Repression of FT to Promote Flowering. *Plant Physiology,* p. pp.00596.02019.

Jones, J., D. G. and Dangl, J., L., 2006. The plant immune system. *Nature,* 444(7117), p. 323.

Jones, J.D.G., Vance, R.E. and Dangl, J.L., 2016. Intracellular innate immune surveillance devices in plants and animals. *Science,* 354(6316).

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R. and Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics (Oxford, England),* 30(9), pp. 1236-1240.

Jouet, A., Saunders, D.G.O., McMullan, M., Ward, B., Furzer, O., Jupe, F., Cevik, V., Hein, I., Thilliez, G.J.A., Holub, E., van Oosterhout, C. and Jones, J.D.G., 2019. Albugo candida race diversity, ploidy and host-associated microbes revealed using DNA sequence capture on diseased plants in the field. *New Phytologist,* 221(3), pp. 1529-1543.

Judelson, H. and Ah-Fong, A.M.V., 2019. Exchanges at the plant-oomycete interface that influence disease. *Plant Physiology,* p. pp.00979.02018.

Järving, R., Lõokene, A., Kurg, R., Siimon, L., Järving, I. and Samel, N., 2012. Activation of 11R-Lipoxygenase Is Fully Ca2+-Dependent and Controlled by the Phospholipid Composition of the Target Membrane. *Biochemistry,* 51(15), pp. 3310-3320.

Kabbage, M., Kessens, R., Bartholomay, L.C. and Williams, B., 2017. The Life and Death of a Plant Cell. *Annual Review of Plant Biology,* 68(1), pp. arplant-043015-111655.

Kadota, Y., Shirasu, K. and Zipfel, C., 2015. Regulation of the NADPH Oxidase RBOHD During Plant Immunity. *Plant and Cell Physiology,* 56(8), pp. 1472-1480.

Kadota, Y., Sklenar, J., Derbyshire, P., Stransfeld, L., Asai, S., Ntoukakis, V., Jones, Jonathan D., Shirasu, K., Menke, F., Jones, A. and Zipfel, C., 2014. Direct Regulation of the NADPH Oxidase RBOHD by the PRR-Associated Kinase BIK1 during Plant Immunity. *Molecular Cell,* 54(1), pp. 43-55.

Keinath, N.F., Waadt, R., Brugman, R., Schroeder, Julian I., Grossmann, G., Schumacher, K. and Krebs, M., 2015. Live Cell Imaging with R-GECO1 Sheds Light on flg22- and Chitin-Induced Transient [Ca2+]cyt Patterns in Arabidopsis. *Molecular Plant,* 8(8), pp. 1188-1200.

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., Baranasic, D., Arenillas, D.J., Sandelin, A., Vandepoele, K., Lenhard, B., Ballester, B., Wasserman, W.W., Parcy, F. and Mathelier, A., 2018. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic acids research,* 46(D1), pp. D260-D266.

Kidokoro, S., Yoneda, K., Takasaki, H., Takahashi, F., Shinozaki, K. and Yamaguchi-Shinozaki, K., 2017. Different Cold-Signaling Pathways Function in the Responses to Rapid and Gradual Decreases in Temperature. *The Plant Cell,* 29(4), pp. 760-774.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology,* 14(4), p. R36.

Kim, Y., Park, S., Gilmour, S.J. and Thomashow, M.F., 2013. Roles of CAMTA transcription factors and salicylic acid in configuring the low-temperature transcriptome and freezing tolerance of Arabidopsis. *The Plant Journal,* 75(3), pp. 364-376.

Kim, Y.S., An, C., Park, S., Gilmour, S.J., Wang, L., Renna, L., Brandizzi, F., Grumet, R. and Thomashow, M.F., 2017. CAMTA-Mediated Regulation of Salicylic Acid Immunity Pathway Genes in Arabidopsis Exposed to Low Temperature and Pathogen Infection. *The Plant Cell,* 29(10), pp. 2465-2477.

King, S.R.F., McLellan, H., Boevink, P.C., Armstrong, M.R., Bukharova, T., Sukarta, O., Win, J., Kamoun, S., Birch, P.R.J. and Banfield, M.J., 2014. *Phytophthora infestans* RXLR Effector PexRD2 Interacts with Host MAPKKKε to Suppress Plant Immune Signaling. *The Plant Cell,* 26(3), pp. 1345-1359.

Knepper, C., Savory, E.A. and Day, B., 2011. The role of NDR1 in pathogen perception and plant defense signaling. *Plant signaling & behavior,* 6(8), pp. 1114-1116.

Knizewski, L., Ginalski, K. and Jerzmanowski, A., 2008. Snf2 proteins in plants: gene silencing and beyond. *Trends in Plant Science,* 13(10), pp. 557-565.

Kobayashi, M., Ohura, I., Kawakita, K., Yokota, N., Fujiwara, M., Shimamoto, K., Doke, N. and Yoshioka, H., 2007. Calcium-Dependent Protein Kinases Regulate the Production of Reactive Oxygen Species by Potato NADPH Oxidase. *The Plant Cell,* 19(3), p. 1065.

Kosugi S., Hasebe M., Tomita M. and H, Y., 2009. cNLS Mapper. Available from: *http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi* [Accessed 2020].

Kosuta, S., Hazledine, S., Sun, J., Miwa, H., Morris, R., Downie, J. and Oldroyd, G., 2008. Differential and chaotic calcium signatures in the symbiosis signaling pathway of legumes. *Proceedings Of The National Academy Of Sciences Of The United States Of Ame,* 105(28), pp. 9823-9828.

Kourelis, J., van der Hoorn, R.A.L. and Sueldo, D.J., 2016. Decoy Engineering The Next Step in Resistance Breeding. *Trends in Plant Science,* 21(5), pp. 371-373.

Kroj, T., Chanclud, E., Michel-Romiti, C., Grand, X. and Morel, J.-B., 2016. Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *New Phytologist,* 210(2), pp. 618-626.

Kudla, J., Batistič, O. and Hashimoto, K., 2010. Calcium Signals: The Lead Currency of Plant Information Processing. *The Plant Cell,* 22(3), pp. 541-563.

Kudla, J., Becker, D., Grill, E., Hedrich, R., Hippler, M., Kummer, U., Parniske, M., Romeis, T. and Schumacher, K., 2018. Advances and current challenges in calcium signaling. *New Phytologist,* 218(2), pp. 414-431.

Kusajima, M., Okumura, Y., Fujita, M. and Nakashita, H., 2017. Abscisic acid modulates salicylic acid biosynthesis for systemic acquired resistance in tomato. *Bioscience, Biotechnology, and Biochemistry,* 81(9), pp. 1850-1853.

La Verde, V., Dominici, P. and Astegno, A., 2018. Towards Understanding Plant Calcium Signaling through Calmodulin-Like Proteins: A Biochemical and Structural Perspective. *International journal of molecular sciences,* 19(5), p. 1331.

Lai, Y. and Eulgem, T., 2018. Transcript-level expression control of plant NLR genes. *Molecular Plant Pathology,* 19(5), pp. 1267-1281.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., Lee, C.H., Nelson,

W.D., Ploetz, L., Singh, S., Wensel, A. and Huala, E., 2011. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research,* 40(D1), pp. D1202-D1210.

Lapin, D., Kovacova, V., Sun, X., Dongus, J.A., Bhandari, D.D., von Born, P., Bautor, J., Guarneri, N., Rzemieniewski, J., Stuttmann, J., Beyer, A. and Parker, J.E., 2019. A coevolved EDS1-SAG101-NRG1 module mediates cell death signaling by TIR-domain immune receptors. *The Plant Cell,* p. tpc.00118.02019.

Lau, N.-C., Kolkman, A., van Schaik, Frederik M.A., Mulder, Klaas W., Pijnappel, W.W.M.P., Heck, Albert J.R. and Timmers, H.T.M., 2009. Human Ccr4–Not complexes contain variable deadenylase subunits. *Biochemical Journal,* 422(3), pp. 443-453.

Le Roux, C., Huet, G., Jauneau, A., Camborde, L., Tremousaygue, D., Kraut, A., Zhou, B., Levaillant, M., Adachi, H., Yoshioka, H., Raffaele, S., Berthome, R., Coute, Y., Parker, J.E. and Deslandes, L., 2015. A Receptor Pair with an Integrated Decoy Converts Pathogen Disabling of Transcription Factors to Immunity. *Cell,* 161(5), pp. 1074-1088.

Le, S.Q. and Gascuel, O., 2008. An Improved General Amino Acid Replacement Matrix. *Molecular Biology and Evolution,* 25(7), pp. 1307-1320.

Leach, J.E., Triplett, L.R., Argueso, C.T. and Trivedi, P., 2017. Communication in the Phytobiome. *Cell,* 169(4), pp. 587-596.

Lee, S.-j., Lee, M.H., Kim, J.-I. and Kim, S.Y., 2014. Arabidopsis Putative MAP Kinase Kinase Kinases Raf10 and Raf11 are Positive Regulators of Seed Dormancy and ABA Response. *Plant and Cell Physiology,* 56(1), pp. 84-97.

Leipe, D.D., Koonin, E.V. and Aravind, L., 2004. STAND, a Class of P-Loop NTPases Including Animal and Plant Regulators of Programmed Cell Death: Multiple, Complex Domain Architectures, Unusual Phyletic Patterns, and Evolution by Horizontal Gene Transfer. *Journal of Molecular Biology,* 343(1), pp. 1-28.

Letunic, I. and Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res,* 44(W1), pp. W242-245.

Letunic, I., Doerks, T. and Bork, P., 2015. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research,* 43(D1), pp. D257-D260.

Li, C., Gu, L., Gao, L., Chen, C., Wei, C.-Q., Qiu, Q., Chien, C.-W., Wang, S., Jiang, L., Ai, L.-F., Chen, C.-Y., Yang, S., Nguyen, V., Qi, Y., Snyder, M.P., Burlingame, A.L., Kohalmi, S.E., Huang, S., Cao, X., Wang, Z.-Y., Wu, K., Chen, X. and Cui, Y., 2016. Concerted genomic targeting of H3K27 demethylase REF6 and chromatin-remodeling ATPase BRM in Arabidopsis. *Nature Genetics,* 48(6), pp. 687-693.

Li, L., Kim, P., Yu, L., Cai, G., Chen, S., Alfano, James R. and Zhou, J.-M., 2016. Activation-Dependent Destruction of a Co-receptor by a Pseudomonas syringae Effector Dampens Plant Immunity. *Cell Host & Microbe,* 20(4), pp. 504-514.

Li, L., Li, M., Yu, L., Zhou, Z., Liang, X., Liu, Z., Cai, G., Gao, L., Zhang, X., Wang, Y., Chen, S. and Zhou, J.-M., 2014. The FLS2-Associated Kinase BIK1 Directly Phosphorylates the NADPH Oxidase RbohD to Control Plant Immunity. *Cell Host & Microbe,* 15(3), pp. 329-338.

Li, N. and Li, Y., 2016. Signaling pathways of seed size control in plants. *Current Opinion in Plant Biology,* 33, pp. 23-32.

Li, S., Wang, Y., Zhao, Y., Zhao, X., Chen, X. and Gong, Z., 2020. Global Co-transcriptional Splicing in Arabidopsis and the Correlation with Splicing Regulation in Mature RNAs. *Molecular Plant,* 13(2), pp. 266-277.

Li, X., Clarke, J.D., Zhang, Y. and Dong, X., 2001. Activation of an EDS1-Mediated R-Gene Pathway in the snc1 Mutant Leads to Constitutive, NPR1-Independent Pathogen Resistance. *Molecular Plant-Microbe Interactions,* 14(10), pp. 1131-1139.

Li, Y. and Li, H., 2012. Many keys to push: diversifying the 'readership' of plant homeodomain fingers. *Acta Biochimica et Biophysica Sinica,* 44(1), pp. 28-39.

Li, Y., Yang, S., Yang, H. and Hua, J., 2007. The TIR-NB-LRR Gene SNC1 Is Regulated at the Transcript Level by Multiple Factors. *Molecular Plant-Microbe Interactions®,* 20(11), pp. 1449-1456.

Li, Y., Zheng, L., Corke, F., Smith, C. and Bevan, M.W., 2008. Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. *Genes and Development,* 22(10), pp. 1331-1336.

Li, Y., Zhou, M., Hu, Q., Bai, X.-c., Huang, W., Scheres, S.H.W. and Shi, Y., 2017. Mechanistic insights into caspase-9 activation by the structure of the apoptosome holoenzyme. *Proceedings of the National Academy of Sciences,* 114(7), pp. 1542-1547.

Lievens, L., Pollier, J., Goossens, A., Beyaert, R. and Staal, J., 2017. Abscisic Acid as Pathogen Effector and Immune Regulator. *Frontiers in plant science,* 8, pp. 587-587.

Lind, C., Dreyer, I., López-Sanjurjo, Enrique J., von Meyer, K., Ishizaki, K., Kohchi, T., Lang, D., Zhao, Y., Kreuzer, I., Al-Rasheid, Khaled A.S., Ronne, H., Reski, R., Zhu, J.-K., Geiger, D. and Hedrich, R., 2015. Stomatal Guard Cells Co-opted an Ancient ABA-Dependent Desiccation Survival System to Regulate Stomatal Closure. *Current Biology,* 25(7), pp. 928-935.

Liu, J., Yang, H., Bao, F., Ao, K., Zhang, X., Zhang, Y. and Yang, S., 2015. IBR5 Modulates Temperature-Dependent, R Protein CHS3-Mediated Defense Responses in Arabidopsis. *PLOS Genetics,* 11(10), p. e1005584.

Liu, N., Hake, K., Wang, W., Zhao, T., Romeis, T. and Tang, D., 2017. CALCIUM-DEPENDENT PROTEIN KINASE5 associates with the truncated NLR protein TIR-NBS2 to contribute to exo70B1-mediated immunity. *The Plant Cell.*

Liu, S., Kracher, B., Ziegler, J., Birkenbihl, R.P. and Somssich, I.E., 2015. Negative regulation of ABA signaling by WRKY33 is critical for Arabidopsis immunity towards Botrytis cinerea 2100. *eLife,* 4, pp. e07295-e07295.

Loake, G. and Grant, M., 2007. Salicylic acid in plant defence—the players and protagonists. *Current Opinion in Plant Biology,* 10(5), pp. 466-472.

Lolle, S., Greeff, C., Petersen, K., Roux, M., Jensen, M.K., Bressendorff, S., Rodriguez, E., Sømark, K., Mundy, J. and Petersen, M., 2017. Matching NLR Immune Receptors to Autoimmunity in camta3 Mutants Using Antimorphic NLR Alleles. *Cell Host & Microbe,* 21(4), pp. 518-529.e514.

Lolle, S., Stevens, D. and Coaker, G., 2020. Plant NLR-triggered immunity: from receptor activation to downstream signaling. *Current Opinion in Immunology,* 62, pp. 99-105.

Love, M.I., Huber, W. and Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology,* 15(12), p. 550.

Lu, D., Wu, S., Gao, X., Zhang, Y., Shan, L. and He, P., 2010. A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proceedings of the National Academy of Sciences,* 107(1), pp. 496-501.

Lumba, S., Toh, S., Handfield, L.-F., Swan, M., Liu, R., Youn, J.-Y., Cutler, Sean R., Subramaniam, R., Provart, N., Moses, A., Desveaux, D. and McCourt, P., 2014. A Mesoscale Abscisic Acid Hormone Interactome Reveals a Dynamic Signaling Landscape in Arabidopsis. *Developmental Cell,* 29(3), pp. 360-372.

Lyons, E. and Freeling, M., 2008. How to usefully compare homologous plant genes and chromosomes as DNA sequences. *The Plant Journal,* 53(4), pp. 661-673.

Ma, Y., Guo, H., Hu, L., Martinez, P.P., Moschou, P.N., Cevik, V., Ding, P., Duxbury, Z., Sarris, P.F. and Jones, J.D.G., 2018. Distinct modes of derepression of an Arabidopsis immune receptor complex by two different bacterial effectors. *Proceedings of the National Academy of Sciences*.

Macho, A.P., Schwessinger, B., Ntoukakis, V., Brutus, A., Segonzac, C., Roy, S., Kadota, Y., Oh, M.-H., Sklenar, J., Derbyshire, P., Lozano-Durán, R., Malinovsky, F.G., Monaghan, J., Menke, F.L., Huber, S.C., He, S.Y. and Zipfel, C., 2014. A Bacterial Tyrosine Phosphatase Inhibits Plant Pattern Recognition Receptor Activation. *Science,* 343(6178), p. 1509.

MacLean, A.M., Bravo, A. and Harrison, M.J., 2017. Plant Signaling and Metabolic Pathways Enabling Arbuscular Mycorrhizal Symbiosis. *The Plant Cell,* 29(10), pp. 2319-2335.

Maqbool, A., Saitoh, H., Franceschetti, M., Stevenson, C.E.M., Uemura, A., Kanzaki, H., Kamoun, S., Terauchi, R. and Banfield, M.J., 2015. Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. *eLife,* 4, p. e08709.

March-Díaz, R., García-Domínguez, M., Lozano-Juste, J., León, J., Florencio, F.J. and Reyes, J.C., 2008. Histone H2A.Z and homologues of components of the SWR1 complex are required to control immunity in Arabidopsis. *The Plant Journal,* 53(3), pp. 475-487.

Marfella, C.G.A. and Imbalzano, A.N., 2007. The Chd family of chromatin remodelers. *Mutation research,* 618(1-2), pp. 30-40.

Martín-Rodríguez, J.A., Huertas, R., Ho-Plágaro, T., Ocampo, J.A., Turečková, V., Tarkowská, D., Ludwig-Müller, J. and García-Garrido, J.M., 2016. Gibberellin-Abscisic Acid Balances during Arbuscular Mycorrhiza Formation in Tomato. *Frontiers in plant science,* 7, pp. 1273-1273.

McMullan, M., Gardiner, A., Bailey, K., Kemen, E., Ward, B.J., Cevik, V., Robert-Seilaniantz, A., Schultz-Larsen, T., Balmuth, A., Holub, E., van Oosterhout, C. and Jones, J.D.G., 2015. Evidence for suppression of immunity as a driver for genomic introgressions and host range expansion in races of Albugo candida, a generalist parasite. *eLife,* 4, p. e04550.

McNeece, B.T., Pant, S.R., Sharma, K., Niruala, P., Lawrence, G.W. and Klink, V.P., 2017. A Glycine max homolog of NON-RACE SPECIFIC DISEASE RESISTANCE 1 (NDR1) alters defense gene expression while functioning during a resistance response to different root pathogens in different genetic backgrounds. *Plant Physiology and Biochemistry,* 114, pp. 60-71.

Meena, M.K., Prajapati, R., Krishna, D., Divakaran, K., Pandey, Y., Reichelt, M., Mathew, M.K., Boland, W., Mithöfer, A. and Vadassery, J., 2019. The Ca^{2+} Channel CNGC19 Regulates Arabidopsis Defense Against Spodoptera Herbivory. *The Plant Cell,* 31(7), pp. 1539-1562.

Meguro, A. and Sato, Y., 2014. Salicylic acid antagonizes abscisic acid inhibition of shoot growth and cell cycle progression in rice. *Scientific Reports,* 4(1), p. 4555.

Meng, X. and Zhang, S., 2013. MAPK Cascades in Plant Disease Resistance Signaling. *Annual Review of Phytopathology,* 51(1), pp. 245-266.

Mentlak, T.A., Kombrink, A., Shinya, T., Ryder, L.S., Otomo, I., Saitoh, H., Terauchi, R., Nishizawa, Y., Shibuya, N., Thomma, B.P.H.J. and Talbot, N.J., 2012. Effector-Mediated Suppression of Chitin-Triggered Immunity by *Magnaporthe oryzae* Is Necessary for Rice Blast Disease. *The Plant Cell,* 24(1), pp. 322-335.

Meunier, E. and Broz, P., 2017. Evolutionary Convergence and Divergence in NLR Function and Structure. *Trends in Immunology,* 38(10), pp. 744-757.

Meyers, B.C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R.W., 2003. Genome-Wide Analysis of NBS-LRR-Encoding Genes in Arabidopsis. *The Plant Cell,* 15(4), pp. 809-834.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H. and Shibuya, N., 2007. CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. *Proceedings of the National Academy of Sciences,* 104(49), pp. 19613-19618.

Mochizuki, S. and Matsui, K., 2018. Green leaf volatile-burst in Arabidopsis is governed by galactolipid oxygenation by a lipoxygenase that is under control of calcium ion. *Biochemical and Biophysical Research Communications,* 505(3), pp. 939-944.

Moeder, W., Ung, H., Mosher, S. and Yoshioka, K., 2010. SA-ABA antagonism in defense responses. *Plant Signaling & Behavior,* 5(10), pp. 1231-1233.

Mohr, T.J., Mammarella, N.D., Hoff, T., Woffenden, B.J., Jelesko, J.G. and McDowell, J.M., 2010. The Arabidopsis Downy Mildew Resistance Gene RPP8 Is Induced by Pathogens and Salicylic Acid and Is Regulated by W Box cis Elements. *Molecular Plant-Microbe Interactions®,* 23(10), pp. 1303-1315.

Monaghan, J., Germain, H., Weihmann, T. and Li, X., 2010. Dissecting plant defence signal transduction: modifiers of snc1 in Arabidopsis. *Canadian Journal of Plant Pathology,* 32(1), pp. 35-42.

Myers, N. and Kent, J., 2003. New consumers: The influence of affluence on the environment. *Proceedings of the National Academy of Sciences,* 100(8), pp. 4963-4968.

Nandety, R.S., Caplan, J.L., Cavanaugh, K., Perroud, B., Wroblewski, T., Michelmore, R.W. and Meyers, B.C., 2013. The Role of TIR-NBS and TIR-X Proteins in Plant Basal Defense Responses. *Plant Physiology,* 162(3), pp. 1459-1472.

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M. and Narusaka, Y., 2009. RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *The Plant Journal,* 60(2), pp. 218-226.

Nepal, M., Andersen, E., Neupane, S. and Benson, B., 2017. Comparative Genomics of Non-TNL Disease Resistance Genes from Six Plant Species. *Genes,* 8(10), p. 249.

Newman, T.E., Lee, J., Williams, S.J., Choi, S., Halane, M.K., Zhou, J., Solomon, P., Kobe, B., Jones, J.D.G., Segonzac, C. and Sohn, K.H., 2018. Autoimmunity and effector recognition in Arabidopsis thaliana can be uncoupled by mutations in the RRS1-R immune receptor. *New Phytologist,* 0(ja).

Nguyen, Q.T.C., Lee, S.-J., Choi, S.-W., Na, Y.-J., Song, M.-R., Hoang, Q.T.N., Sim, S.Y., Kim, M.-S., Kim, J.-I., Soh, M.-S. and Kim, S.Y., 2019. Arabidopsis Raf-Like Kinase Raf10 Is a Regulatory Component of Core ABA Signaling. *Molecules and cells,* 42(9), pp. 646-660.

Nour-Eldin, H.H., Geu-Flores, F. and Halkier, B.A., 2010. USER cloning and USER fusion: The ideal cloning techniques for small and big laboratories. *Methods in Molecular Biology,* 643, pp. 185-200.

Nützmann, H.-W. and Osbourn, A., 2015. Regulation of metabolic gene clusters in Arabidopsis thaliana. *New Phytologist,* 205(2), pp. 503-510.

Ohmura, T., Ueda, T., Hashimoto, Y. and Imoto, T., 2001. Tolerance of point substitution of methionine for isoleucine in hen egg white lysozyme. *Protein Engineering, Design and Selection,* 14(6), pp. 421-425.

Ojolo, S.P., Cao, S., Priyadarshani, S.V.G.N., Li, W., Yan, M., Aslam, M., Zhao, H. and Qin, Y., 2018. Regulation of Plant Growth and Development: A Review From a Chromatin Remodeling Perspective. *Frontiers in plant science,* 9, pp. 1232-1232.

Okuyama, Y., Kanzaki, H., Abe, A., Yoshida, K., Tamiru, M., Saitoh, H., Fujibe, T., Matsumura, H., Shenton, M., Galam, D.C., Undan, J., Ito, A., Sone, T. and Terauchi, R., 2011. A multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. *The Plant Journal,* 66(3), pp. 467-479.

Oldroyd, G., E. D., 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology,* 11(4), p. 252.

Ortiz, D., De Guillen, K., Cesari, S., Chalvon, V., Gracy, J., Padilla, A. and Kroj, T., 2017. Recognition of the Magnaporthe oryzae effector AVR-Pia by the decoy domain of the rice NLR immune receptor RGA5. *The Plant Cell*.

Palma, K., Zhao, Q., Cheng, Y.T., Bi, D., Monaghan, J., Cheng, W., Zhang, Y. and Li, X., 2007. Regulation of plant innate immunity by three proteins in a complex conserved across the plant and animal kingdoms. *Genes & Development,* 21(12), pp. 1484-1493.

Pan, Y., Chai, X., Gao, Q., Zhou, L., Zhang, S., Li, L. and Luan, S., 2019. Dynamic Interactions of Plant CNGC Subunits and Calmodulins Drive Oscillatory Ca2+ Channel Activities. *Developmental Cell*, 48(5), pp. 710-725.e715.

Paulus, J.K., Kourelis, J. and van der Hoorn, R.A.L., 2017. Bodyguards: Pathogen-Derived Decoys That Protect Virulence Factors. *Trends in Plant Science*.

Peng, Y., Chen, L., Lu, Y., Wu, Y., Lia, Y., Zhu, Z., Dumenil, J. and Bevan, M.W., 2015. The Ubiquitin receptors DA1, DAR1, and DAR2 redundantly regulate endoreduplication by modulating the stability of TCP14/ 15 in arabidopsis. *Plant Cell*, 27(3), pp. 649-662.

Peng, Y., van Wersch, R. and Zhang, Y., 2017. Convergent and Divergent Signaling in PAMP-Triggered Immunity and Effector-Triggered Immunity. *Molecular Plant-Microbe Interactions®*, 31(4), pp. 403-409.

Peng, Z., Hu, Y., Zhang, J., Huguet-Tapia, J.C., Block, A.K., Park, S., Sapkota, S., Liu, Z., Liu, S. and White, F.F., 2019. *Xanthomonas translucens* commandeers the host rate-limiting step in ABA biosynthesis for disease susceptibility. *Proceedings of the National Academy of Sciences*, 116(42), pp. 20938-20946.

Plett, J.M. and Martin, F.M., 2017. Know your Enemy, Embrace your Friend: Using omics to understand how plants respond differently to pathogenic and mutualistic microorganisms. *The Plant Journal*.

Poovaiah, B.W., Du, L., Wang, H. and Yang, T., 2013. Recent Advances in Calcium/Calmodulin-Mediated Signaling with an Emphasis on Plant-Microbe Interactions. *Plant Physiology*, 163(2), pp. 531-542.

Porebski, S., Bailey, L.G. and Baum, B.R., 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter*, 15(1), pp. 8-15.

Presti, L.L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S. and Kahmann, R., 2015. Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology*, 66(1), pp. 513-545.

Prince, D.C., Rallapalli, G., Xu, D., Schoonbeek, H.-j., Çevik, V., Asai, S., Kemen, E., Cruz-Mireles, N., Kemen, A., Belhaj, K., Schornack, S., Kamoun, S., Holub, E.B., Halkier, B.A. and Jones, J.D.G., 2017. Albugo-imposed changes to tryptophan-derived antimicrobial metabolite biosynthesis may contribute to suppression of non-host resistance to Phytophthora infestans in Arabidopsis thaliana. *BMC Biology*, 15(1), p. 20.

Qi, T., Seong, K., Thomazella, D.P.T., Kim, J.R., Pham, J., Seo, E., Cho, M.-J., Schultink, A. and Staskawicz, B.J., 2018. NRG1 functions downstream of EDS1 to regulate TIR-NLR-mediated plant immunity in *Nicotiana benthamiana*. *Proceedings of the National Academy of Sciences*.

Rahman, H., Yang, J., Xu, Y.-P., Munyampundu, J.-P. and Cai, X.-Z., 2016. Phylogeny of Plant CAMTAs and Role of AtCAMTAs in Nonhost Resistance to Xanthomonas oryzae pv. oryzae. *Frontiers in Plant Science*, 7(177).

Ramirez-Prado, J.S., Piquerez, S.J.M., Bendahmane, A., Hirt, H., Raynaud, C. and Benhamed, M., 2018. Modify the Histone to Win the Battle: Chromatin Dynamics in Plant–Pathogen Interactions. *Frontiers in Plant Science*, 9(355).

Ramón-Maiques, S., Kuo, A.J., Carney, D., Matthews, A.G.W., Oettinger, M.A., Gozani, O. and Yang, W., 2007. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proceedings of the National Academy of Sciences,* 104(48), pp. 18993-18998.

Ranf, S., Eschen-Lippold, L., Pecher, P., Lee, J. and Scheel, D., 2011. Interplay between calcium signalling and early signalling elements during defence responses to microbe- or damage-associated molecular patterns. *The Plant Journal,* 68(1), pp. 100-113.

Ranty, B., Aldon, D. and Galaud, J.-P., 2006. Plant calmodulins and calmodulin-related proteins: multifaceted relays to decode calcium signals. *Plant signaling & behavior,* 1(3), pp. 96-104.

Rodriguez-Moreno, L., Ebert, M.K., Bolton, M.D. and Thomma, B.P.H.J., 2017. Tools of the crook – infection strategies of fungal plant pathogens. *The Plant Journal,* pp. n/a-n/a.

Rodriguez-Moreno, L., Song, Y. and Thomma, B.P.H.J., 2017. Transfer and engineering of immune receptors to improve recognition capacities in crops. *Current Opinion in Plant Biology,* 38, pp. 42-49.

Roux, F., Noël, L., Rivas, S. and Roby, D., 2014. ZRK atypical kinases: emerging signaling components of plant immunity. *New Phytologist,* 203(3), pp. 713-716.

Sah, S.K., Reddy, K.R. and Li, J., 2016. Abscisic Acid and Abiotic Stress Tolerance in Crop Plants. *Frontiers in Plant Science,* 7(571).

Saharan, G.S. and Verma, P.R., 1992. *White Rusts: A review of economically important species.* Ottawa: International Development Research Centre. (IDRC- MR315e: IV+65 p).

Saharan, G.S., Verma, P.R., Meena, P.D. and Kumar, A., 2014. White Rust of Crucifers: Biology, Ecology and Management. In: G.S. Saharan, P.R. Verma, P.D. Meena and A. Kumar, eds. *White Rust of Crucifers: Biology, Ecology and Management.* New Delhi: Springer India, pp. 1-6.

Saijo, Y., Loo, E.P.-i. and Yasuda, S., 2018. Pattern recognition receptors and signaling in plant–microbe interactions. *The Plant Journal,* 93(4), pp. 592-613.

Sanchez, R. and Zhou, M.-M., 2011. The PHD finger: a versatile epigenome reader. *Trends in biochemical sciences,* 36(7), pp. 364-372.

Sang, Q., Pajoro, A., Sun, H., Song, B., Yang, X., Stolze, S.C., Andr**&#65533;**s, F., Schneeberger, K., Nakagami, H. and Coupland, G., 2020. Mutagenesis of a Quintuple Mutant Impaired in Environmental Responses Reveals Roles for CHROMATIN REMODELING4 in the Arabidopsis Floral Transition. *The Plant Cell,* p. tpc.00992.02019.

Sarris, P.F., Cevik, V., Dagdas, G., Jones, J.D.G. and Krasileva, K.V., 2016. Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *Bmc Biology,* 14.

Sarris, P.F., Duxbury, Z., Huh, S.U., Ma, Y., Segonzac, C., Sklenar, J., Derbyshire, P., Cevik, V., Rallapalli, G., Saucet, S.B., Wirthmueller, L., Menke, F.L.H., Sohn, K.H. and Jones, J.D.G., 2015. A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors. *Cell,* 161(5), pp. 1089-1100.

Sasayama, D., Matsuoka, D., Oka, M., Shitamichi, N., Furuya, T., Azuma, T., Itoh, K. and Nanmori, T., 2011. MAP3Kδ4, an *Arabidopsis* Raf-like MAP3K, regulates plant growth and shoot branching. *Plant Biotechnology,* 28(5), pp. 463-470.

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N. and Nelson, A., 2019. The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution,* 3(3), pp. 430-439.

Schmeichel, K.L. and Beckerle, M.C., 1994. The LIM domain is a modular protein-binding interface. *Cell,* 79(2), pp. 211-219.

Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P., 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences,* 95(11), pp. 5857-5864.

Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D. and Yu, T.-H., 2008. Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change. *Science,* 319(5867), pp. 1238-1240.

Shao, Z.-Q., Xue, J.-Y., Wang, Q., Wang, B. and Chen, J.-Q., 2018. Revisiting the Origin of Plant NBS-LRR Genes. *Trends in Plant Science*.

Shao, Z.-Q., Xue, J.-Y., Wu, P., Zhang, Y.-M., Wu, Y., Hang, Y.-Y., Wang, B. and Chen, J.-Q., 2016. Large-Scale Analyses of Angiosperm Nucleotide-Binding Site-Leucine-Rich Repeat Genes Reveal Three Anciently Diverged Classes with Distinct Evolutionary Patterns. *Plant Physiology,* 170(4), pp. 2095-2109.

Shao, Z.-Q., Zhang, Y.-M., Hang, Y.-Y., Xue, J.-Y., Zhou, G.-C., Wu, P., Wu, X.-Y., Wu, X.-Z., Wang, Q., Wang, B. and Chen, J.-Q., 2014. Long-Term Evolution of Nucleotide-Binding Site-Leucine-Rich Repeat Genes: Understanding Gained from and beyond the Legume Family. *Plant Physiology,* 166(1), pp. 217-234.

Shimada, T.L., Shimada, T. and Hara-Nishimura, I., 2010. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. *The Plant Journal,* 61(3), pp. 519-528.

Shine, M.B., Yang, J.W., El-Habbak, M., Nagyabhyru, P., Fu, D.Q., Navarre, D., Ghabrial, S., Kachroo, P. and Kachroo, A., 2016. Cooperative functioning between phenylalanine ammonia lyase and isochorismate synthase activities contributes to salicylic acid biosynthesis in soybean. *New Phytologist*.

Shitamichi, N., Matsuoka, D., Sasayama, D., Furuya, T. and Nanmori, T., 2013. Over-expression of MAP3Kδ4, an ABA-inducible Raf-like MAP3K that confers salt tolerance in *Arabidopsis*. *Plant Biotechnology,* 30(2), pp. 111-118.

Sikora, P., Chawade, A., Larsson, M., Olsson, J. and Olsson, O., 2012. Mutagenesis as a Tool in Plant Genetics, Functional Genomics, and Breeding. *International Journal of Plant Genomics,* 2011.

Sims, R.J., Chen, C.-F., Santos-Rosa, H., Kouzarides, T., Patel, S.S. and Reinberg, D., 2005. Human but Not Yeast CHD1 Binds Directly and Selectively to Histone H3 Methylated at Lysine 4 via Its Tandem Chromodomains. *Journal of Biological Chemistry,* 280(51), pp. 41789-41792.

Singh, A., Guest, D. and Copeland, L., 2015. Associations Between Glucosinolates, White Rust, and Plant Defense Activators in Brassica Plants: A Review. *International Journal of Vegetable Science*, 21(3), pp. 297-313.

Slootweg, E., Roosien, J., Spiridon, L.N., Petrescu, A.-J., Tameling, W., Joosten, M., Pomp, R., van Schaik, C., Dees, R., Borst, J.W., Smant, G., Schots, A., Bakker, J. and Goverse, A., 2010. Nucleocytoplasmic Distribution Is Required for Activation of Resistance by the Potato NB-LRR Receptor Rx1 and Is Balanced by Its Functional Domains. *The Plant Cell*, 22(12), pp. 4195-4215.

Smakowska-Luzan, E., Mott, G.A., Parys, K., Stegmann, M., Howton, T.C., Layeghifard, M., Neuhold, J., Lehner, A., Kong, J., Grünwald, K., Weinberger, N., Satbhai, S.B., Mayer, D., Busch, W., Madalinski, M., Stolt-Bergner, P., Provart, N.J., Mukhtar, M.S., Zipfel, C., Desveaux, D., Guttman, D.S. and Belkhadir, Y., 2018. An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. *Nature*, 553, p. 342.

Sohn, K.H., Segonzac, C., Rallapalli, G., Sarris, P.F., Woo, J.Y., Williams, S.J., Newman, T.E., Paek, K.H., Kobe, B. and Jones, J.D.G., 2014. The Nuclear Immune Receptor RPS4 Is Required for RRS1SLH1-Dependent Constitutive Defense Activation in Arabidopsis thaliana. *PLOS Genetics*, 10(10), p. e1004655.

Song, X.-P., Hansen, M.C., Stehman, S.V., Potapov, P.V., Tyukavina, A., Vermote, E.F. and Townshend, J.R., 2018. Global land change from 1982 to 2016. *Nature*, 560(7720), pp. 639-643.

Srivastava, V. and Verma, P.K., 2015. Genome Wide Identification of LIM Genes in Cicer arietinum and Response of Ca- 2LIMs in Development, Hormone and Pathogenic Stress. *PLoS ONE*, 10, p. e0141900.

Srivastava, V. and Verma, P.K., 2017. The plant LIM proteins: unlocking the hidden attractions. *Planta*.

Staiger, D., Korneli, C., Lummer, M. and Navarro, L., 2013. Emerging role for RNA-based regulation in plant immunity. *New Phytologist*, 197(2), pp. 394-404.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp. 1312-1313.

Stein, J.C., Yu, Y., Copetti, D., Zwickl, D.J., Zhang, L., Zhang, C., Chougule, K., Gao, D., Iwata, A., Goicoechea, J.L., Wei, S., Wang, J., Liao, Y., Wang, M., Jacquemin, J., Becker, C., Kudrna, D., Zhang, J., Londono, C.E.M., Song, X., Lee, S., Sanchez, P., Zuccolo, A., Ammiraju, J.S.S., Talag, J., Danowitz, A., Rivera, L.F., Gschwend, A.R., Noutsos, C., Wu, C.-c., Kao, S.-m., Zeng, J.-w., Wei, F.-j., Zhao, Q., Feng, Q., El Baidouri, M., Carpentier, M.-C., Lasserre, E., Cooke, R., Rosa Farias, D.d., da Maia, L.C., dos Santos, R.S., Nyberg, K.G., McNally, K.L., Mauleon, R., Alexandrov, N., Schmutz, J., Flowers, D., Fan, C., Weigel, D., Jena, K.K., Wicker, T., Chen, M., Han, B., Henry, R., Hsing, Y.-i.C., Kurata, N., de Oliveira, A.C., Panaud, O., Jackson, S.A., Machado, C.A., Sanderson, M.J., Long, M., Ware, D. and Wing, R.A., 2018. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. *Nature Genetics*, 50(2), pp. 285-296.

Sukul, P. and Spiteller, M., 2000. Metalaxyl: persistence, degradation, metabolism, and analytical methods. *Rev Environ Contam Toxicol*, 164, pp. 1-26.

Sun, T., Huang, J., Xu, Y., Verma, V., Jing, B., Sun, Y., Ruiz Orduna, A., Tian, H., Huang, X., Xia, S., Schafer, L., Jetter, R., Zhang, Y. and Li, X., 2020. Redundant CAMTA Transcription Factors Negatively Regulate the Biosynthesis of Salicylic Acid and N-Hydroxypipecolic Acid by Modulating the Expression of SARD1 and CBP60g. *Molecular Plant*, 13(1), pp. 144-156.

Sánchez-Vallet, A., Mesters, J.R. and Thomma, B.P.H.J., 2015. The battle for chitin recognition in plant-microbe interactions. *FEMS Microbiology Reviews*, 39(2), pp. 171-183.

Tamborski, J. and Krasileva, K.V., 2020. Evolution of Plant NLRs: From Natural History to Precise Modifications. *Annual Review of Plant Biology*.

Tameling, W.I.L., Vossen, J.H., Albrecht, M., Lengauer, T., Berden, J.A., Haring, M.A., Cornelissen, B.J.C. and Takken, F.L.W., 2006. Mutations in the NB-ARC Domain of I-2 That Impair ATP Hydrolysis Cause Autoactivation. *Plant Physiology*, 140(4), pp. 1233-1245.

Thouly, C., Le Masson, M., Lai, X., Carles, C.C. and Vachon, G., 2020. Unwinding BRAHMA Functions in Plants. *Genes*, 11(1), p. 90.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W. and Su, Z., 2017. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Research*, 45(W1), pp. W122-W129.

Tian, W., Hou, C., Ren, Z., Wang, C., Zhao, F., Dahlbeck, D., Hu, S., Zhang, L., Niu, Q., Li, L., Staskawicz, B.J. and Luan, S., 2019. A calmodulin-gated calcium channel links pathogen patterns to plant immunity. *Nature*.

Ting, J.P.Y., Lovering, R.C., Alnemri, E.S., Bertin, J., Boss, J.M., Davis, B.K., Flavell, R.A., Girardin, S.E., Godzik, A., Harton, J.A., Hoffman, H.M., Hugot, J.-P., Inohara, N., MacKenzie, A., Maltais, L.J., Nunez, G., Ogura, Y., Otten, L.A., Philpott, D., Reed, J.C., Reith, W., Schreiber, S., Steimle, V. and Ward, P.A., 2008. The NLR Gene Family: A Standard Nomenclature. *Immunity*, 28(3), pp. 285-287.

Tokizawa, M., Kobayashi, Y., Saito, T., Kobayashi, M., Iuchi, S., Nomoto, M., Tada, Y., Yamamoto, Y.Y. and Koyama, H., 2015. SENSITIVE TO PROTON RHIZOTOXICITY1, CALMODULIN BINDING TRANSCRIPTION ACTIVATOR2, and Other Transcription Factors Are Involved in *ALUMINUM-ACTIVATED MALATE TRANSPORTER1* Expression. *Plant Physiology*, 167(3), pp. 991-1003.

Toruño, T.Y., Stergiopoulos, I. and Coaker, G., 2016. Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. *Annual Review of Phytopathology*, 54(1), pp. 419-441.

Tsuda, K., Mine, A., Bethke, G., Igarashi, D., Botanga, C.J., Tsuda, Y., Glazebrook, J., Sato, M. and Katagiri, F., 2013. Dual Regulation of Gene Expression Mediated by Extended MAPK Activation and Salicylic Acid Contributes to Robust Innate Immunity in Arabidopsis thaliana. *PLOS Genetics*, 9(12), p. e1004015.

Tuteja, N. and Sopory, S.K., 2008. Plant signaling in stress: G-protein coupled receptors, heterotrimeric G-proteins and signal coupling via phospholipases. *Plant signaling & behavior*, 3(2), pp. 79-86.

Tyagi, M., Imam, N., Verma, K. and Patel, A.K., 2016. Chromatin remodelers: We are the drivers!! *Nucleus (Austin, Tex.)*, 7(4), pp. 388-404.

Uehling, J., Deveau, A. and Paoletti, M., 2017. Do fungi have an innate immune response? An NLR-based comparison to plant and animal immune systems. *PLOS Pathogens,* 13(10), p. e1006578.

United Nations, D.o.E.a.S.A., 2019. *World Population Prospects 2019.* United Nations, Department of Economic and Social Affairs, Population Division.

Urbach, J.M. and Ausubel, F.M., 2017. The NBS-LRR architectures of plant R-proteins and metazoan NLRs evolved in independent events. *Proceedings of the National Academy of Sciences,* 114(5), pp. 1063-1068.

Vadassery, J., Reichelt, M., Hause, B., Gershenzon, J., Boland, W. and Mithöfer, A., 2012. CML42-Mediated Calcium Signaling Coordinates Responses to Spodoptera Herbivory and Abiotic Stresses in Arabidopsis. *Plant Physiology,* 159(3), pp. 1159-1175.

Van de Weyer, A.-L., Monteiro, F., Furzer, O.J., Nishimura, M.T., Cevik, V., Witek, K., Jones, J.D.G., Dangl, J.L., Weigel, D. and Bemm, F., 2019. A Species-Wide Inventory of NLR Genes and Alleles in *Arabidopsis thaliana*. *Cell,* 178(5), pp. 1260-1272.e1214.

Van Der Hoorn, R.A.L. and Kamoun, S., 2008. From guard to decoy: A new model for perception of plant pathogen effectors. *Plant Cell,* 20(8), pp. 2009-2017.

van Doorn, W.G., 2011. Classes of programmed cell death in plants, compared to those in animals. *Journal of Experimental Botany,* 62(14), pp. 4749-4761.

van Wersch, S. and Li, X., 2019. Stronger When Together: Clustering of Plant NLR Disease resistance Genes. *Trends in Plant Science,* 24(8), pp. 688-699.

Vanhaeren, H., Nam, Y.-J., De Milde, L., Chae, E., Storme, V., Weigel, D., Gonzalez, N. and Inzé, D., 2017. Forever Young: The Role of Ubiquitin Receptor DA1 and E3 Ligase BIG BROTHER in Controlling Leaf Growth and Development. *Plant Physiology,* 173(2), pp. 1269-1282.

Vlot, A.C., Dempsey, D.M.A. and Klessig, D.F., 2009. Salicylic Acid, a Multifaceted Hormone to Combat Disease. *Annual Review of Phytopathology,* 47(1), pp. 177-206.

Vogt, J.H.M. and Schippers, J.H.M., 2015. Setting the PAS, the role of circadian PAS domain proteins during environmental adaptation in plants. *Frontiers in plant science,* 6, pp. 513-513.

Wachsman, G., Modliszewski, J.L., Valdes, M. and Benfey, P.N., 2017. A SIMPLE pipeline for mapping point mutations. *Plant Physiology*.

Wagner, S., Stuttmann, J., Rietz, S., Guerois, R., Brunstein, E., Bautor, J., Niefind, K. and Parker, Jane E., 2013. Structural Basis for Signaling by Exclusive EDS1 Heteromeric Complexes with SAG101 or PAD4 in Plant Innate Immunity. *Cell Host & Microbe,* 14(6), pp. 619-630.

Walerowski, P., Gündel, A., Yahaya, N., Truman, W., Sobczak, M., Olszak, M., Rolfe, S.A., Borisjuk, L. and Malinowski, R., 2018. Clubroot Disease Stimulates Early Steps of Phloem Differentiation and Recruits SWEET Sucrose Transporters within Developing Galls. *The Plant Cell*.

Walther, M., Wiesner, R. and Kuhn, H., 2004. Investigations into calcium-dependent membrane association of 15-lipoxygenase-1. Mechanistic roles of surface-exposed hydrophobic amino acids and calcium. *J Biol Chem,* 279(5), pp. 3717-3725.

Wan, L., Essuman, K., Anderson, R.G., Sasaki, Y., Monteiro, F., Chung, E.-H., Osborne Nishimura, E., DiAntonio, A., Milbrandt, J., Dangl, J.L. and Nishimura, M.T., 2019. TIR domains of plant immune receptors are NAD^+-cleaving enzymes that promote cell death. *Science,* 365(6455), pp. 799-803.

Wang, G., Roux, B., Feng, F., Guy, E., Li, L., Li, N., Zhang, X., Lautier, M., Jardinaud, M.-F., Chabannes, M., Arlat, M., Chen, S., He, C., Noel, L.D. and Zhou, J.-M., 2015. The Decoy Substrate of a Pathogen Effector and a Pseudokinase Specify Pathogen-Induced Modified-Self Recognition and Immunity in Plants. *Cell Host & Microbe,* 18(3), pp. 285-295.

Wang, J., Hu, M., Wang, J., Qi, J., Han, Z., Wang, G., Qi, Y., Wang, H.-W., Zhou, J.-M. and Chai, J., 2019. Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science,* 364(6435), p. eaav5870.

Wang, J., Wang, J., Hu, M., Wu, S., Qi, J., Wang, G., Han, Z., Qi, Y., Gao, N., Wang, H.-W., Zhou, J.-M. and Chai, J., 2019. Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science,* 364(6435), p. eaav5868.

Wang, J.-L., Tang, M.-Q., Chen, S., Zheng, X.-F., Mo, H.-X., Li, S.-J., Wang, Z., Zhu, K.-M., Ding, L.-N., Liu, S.-Y., Li, Y.-H. and Tan, X.-L., 2017. Down-regulation of BnDA1, whose gene locus is associated with the seeds weight, improves the seeds weight and organ size in Brassica napus. *Plant Biotechnology Journal,* pp. n/a-n/a.

Wang, L., Tsuda, K., Sato, M., Cohen, J.D., Katagiri, F. and Glazebrook, J., 2009. Arabidopsis CaM Binding Protein CBP60g Contributes to MAMP-Induced SA Accumulation and Is Involved in Disease Resistance against Pseudomonas syringae. *PLOS Pathogens,* 5(2), p. e1000301.

Wang, W., Feng, B., Zhou, J.-M. and Tang, D., 2020. Plant immune signaling: Advancing on two frontiers. *Journal of Integrative Plant Biology,* 62(1), pp. 2-24.

Wang, Y., Tyler, B.M. and Wang, Y., 2019. Defense and Counterdefense During Plant-Pathogenic Oomycete Infection. *Annual Review of Microbiology,* 73(1), pp. 667-696.

Wang, Y., Zhang, Y., Wang, Z., Zhang, X. and Yang, S., 2013. A missense mutation in CHS1, a TIR- NB protein, induces chilling sensitivity in Arabidopsis. *The Plant journal : for cell and molecular biology,* 75(4), p. 553.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J., 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25(9), pp. 1189-1191.

Wiermer, M., Feys, B.J. and Parker, J.E., 2005. Plant immunity: the EDS1 regulatory node. *Current Opinion in Plant Biology,* 8(4), pp. 383-389.

Wildermuth, M., C., Dewdney, J., Wu, G. and Ausubel, F., M., 2001. Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature,* 414(6863), p. 562.

Wildermuth, M.C., Steinwand, M.A., McRae, A.G., Jaenisch, J. and Chandran, D., 2017. Adapted Biotroph Manipulation of Plant Cell Ploidy. *Annual Review of Phytopathology,* 55.

Williams, S., YIn, L., Foley, G., Casey, L., Outram, M., Ericsson, D., Lu, J., Boden, M., Dry, I. and Kobe, B., 2016. Structure and function of the TIR domain from the grape NLR protein RPV1. *Frontiers in Plant Science*, 7(1850).

Williams, S.J., 2014. Structural Basis for Assembly and Function of a Heterodimeric Plant Immune Receptor. *Science*, 344(6181), pp. 299-304.

Wróblewski, T., Spiridon, L., Martin, E.C., Petrescu, A.-J., Cavanaugh, K., Truco, M.J., Xu, H., Gozdowski, D., Pawłowski, K., Michelmore, R.W. and Takken, F.L.W., 2018. Genome-wide functional analyses of plant coiled–coil NLR-type pathogen receptors reveal essential roles of their N-terminal domain in oligomerization, networking, and immunity. *PLOS Biology*, 16(12), p. e2005821.

Wu, C.-H., Abd-El-Haliem, A., Bozkurt, T.O., Belhaj, K., Terauchi, R., Vossen, J.H. and Kamoun, S., 2017. NLR network mediates immunity to diverse plant pathogens. *Proceedings of the National Academy of Sciences*, 114(30), pp. 8113-8118.

Wu, Z., Li, M., Dong, O.X., Xia, S., Liang, W., Bao, Y., Wasteneys, G. and Li, X., 2018. Differential regulation of TNL-mediated immune signaling by redundant helper CNLs. *New Phytologist*, 0(ja).

Xia, S., Cheng, Y.T., Huang, S., Win, J., Soards, A., Jinn, T.-L., Jones, J.D.G., Kamoun, S., Chen, S., Zhang, Y. and Li, X., 2013. Regulation of Transcription of Nucleotide-Binding Leucine-Rich Repeat-Encoding Genes *SNC1* and *RPP4* via H3K4 Trimethylation. *Plant Physiology*, 162(3), pp. 1694-1705.

Xia, T., Li, N., Dumenil, J., Li, J., Kamenski, A., Bevan, M.W., Gao, F. and Li, Y., 2013. The Ubiquitin Receptor DA1 Interacts with the E3 Ubiquitin Ligase DA2 to Regulate Seed and Organ Size in Arabidopsis. *The Plant Cell*, 25(9), pp. 3347-3359.

Xu, F., Xu, S., Wiermer, M., Zhang, Y. and Li, X., 2012a. The cyclin L homolog MOS12 and the MOS4-associated complex are required for the proper splicing of plant resistance genes. *The Plant Journal*, 70(6), pp. 916-928.

Xu, F., Xu, S., Wiermer, M., Zhang, Y. and Li, X., 2012b. The cyclin L homolog MOS12 and the MOS4-associated complex are required for the proper splicing of plant resistance genes. *The Plant Journal*, 70(6), pp. 916-928.

Xu, F., Zhu, C., Cevik, V., Johnson, K., Liu, Y., Sohn, K., Jones, J., Holub, E., B. and Li, X., 2015. Autoimmunity conferred by chs3-2D relies on CSA1, its adjacent TNL-encoding neighbour. *Scientific Reports*, 5.

Yamasaki, K., Motomura, Y., Yagi, Y., Nomura, H., Kikuchi, S., Nakai, M. and Shiina, T., 2013. Chloroplast envelope localization of EDS5, an essential factor for salicylic acid biosynthesis in Arabidopsis thaliana. *Plant signaling & behavior*, 8(4), pp. e23603-e23603.

Yang, H., Shi, Y., Liu, J., Guo, L., Zhang, X. and Yang, S., 2010. A mutant CHS3 protein with TIR- NB- LRR- LIM domains modulates growth, cell death and freezing tolerance in a temperature- dependent manner in Arabidopsis. *Plant Journal*, 63(2), pp. 283-296.

Yasuda, M., Ishikawa, A., Jikumaru, Y., Seki, M., Umezawa, T., Asami, T., Maruyama-Nakashita, A., Kudo, T., Shinozaki, K., Yoshida, S. and Nakashita, H., 2008. Antagonistic interaction between systemic acquired resistance and the abscisic acid-mediated abiotic stress response in Arabidopsis. *The Plant cell*, 20(6), pp. 1678-1692.

Yi, H. and Richards, E.J., 2007. A cluster of disease resistance genes in Arabidopsis is coordinately regulated by transcriptional activation and RNA silencing. *The Plant cell*, 19(9), pp. 2929-2939.

Yip Delormel, T. and Boudsocq, M., 2019. Properties and functions of calcium-dependent protein kinases and their relatives in Arabidopsis thaliana. *New Phytologist*, 0(0).

Yu, N., Nützmann, H.-W., MacDonald, J.T., Moore, B., Field, B., Berriri, S., Trick, M., Rosser, S.J., Kumar, S.V., Freemont, P.S. and Osbourn, A., 2016. Delineation of metabolic gene clusters in plant genomes by chromatin signatures. *Nucleic Acids Research*, 44(5), pp. 2255-2265.

Yu, X., Xu, G., Li, B., de Souza Vespoli, L., Liu, H., Moeder, W., Chen, S., de Oliveira, M.V.V., Ariádina de Souza, S., Shao, W., Rodrigues, B., Ma, Y., Chhajed, S., Xue, S., Berkowitz, G.A., Yoshioka, K., He, P. and Shan, L., 2019. The Receptor Kinases BAK1/SERK4 Regulate Ca2+ Channel-Mediated Cellular Homeostasis for Cell Death Containment. *Current Biology*, 29(22), pp. 3778-3790.e3778.

Yuan, P., Du, L. and Poovaiah, B., 2018. Ca2+/Calmodulin-Dependent AtSR1/CAMTA3 Plays Critical Roles in Balancing Plant Growth and Immunity. *International Journal of Molecular Sciences*, 19(6), p. 1764.

Yuan, P., Jauregui, E., Du, L., Tanaka, K. and Poovaiah, B.W., 2017. Calcium signatures and signaling events orchestrate plant–microbe interactions. *Current Opinion in Plant Biology*, 38, pp. 173-183.

Yue, J.-X., Meyers, B.C., Chen, J.-Q., Tian, D. and Yang, S., 2012. Tracing the origin and evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes. *New Phytologist*, 193(4), pp. 1049-1063.

Zambelli, F., Pesole, G. and Pavesi, G., 2009. Pscan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. *Nucleic acids research*, 37(Web Server issue), pp. W247-W252.

Zander, M., Willige, B.C., He, Y., Nguyen, T.A., Langford, A.E., Nehring, R., Howell, E., McGrath, R., Bartlett, A., Castanon, R., Nery, J.R., Chen, H., Zhang, Z., Jupe, F., Stepanova, A., Schmitz, R.J., Lewsey, M.G., Chory, J. and Ecker, J.R., 2019. Epigenetic silencing of a multifunctional plant stress regulator. *eLife*, 8, p. e47835.

Zbierzak, A.M., Porfirova, S., Griebel, T., Melzer, M., Parker, J.E. and Dörmann, P., 2013. A TIR- NBS protein encoded by Arabidopsis Chilling Sensitive 1 (CHS1) limits chloroplast damage and cell death at low temperature. *The Plant journal : for cell and molecular biology*, 75(4), p. 539.

Zeng, Z., Zhang, W., Marand, A.P., Zhu, B., Buell, C.R. and Jiang, J., 2019. Cold stress induces enhanced chromatin accessibility and bivalent histone modifications H3K4me3 and H3K27me3 of active genes in potato. *Genome Biology*, 20(1), p. 123.

Zhan, J. and McDonald, B.A., 2013. Experimental Measures of Pathogen Competition and Relative Fitness. *Annual Review of Phytopathology*, 51(1), pp. 131-153.

Zhang, H., Bishop, B., Ringenberg, W., Muir, W.M. and Ogas, J., 2012. The CHD3 Remodeler PICKLE Associates with Genes Enriched for Trimethylation of Histone H3 Lysine 27. *Plant Physiology*, 159(1), pp. 418-432.

Zhang, J., Li, W., Xiang, T., Liu, Z., Laluk, K., Ding, X., Zou, Y., Gao, M., Zhang, X., Chen, S., Mengiste, T., Zhang, Y. and Zhou, J.-M., 2010. Receptor-like Cytoplasmic Kinases Integrate Signaling from Multiple Plant Immune Receptors and Are Targeted by a Pseudomonas syringae Effector. *Cell Host & Microbe,* 7(4), pp. 290-301.

Zhang, M. and Coaker, G., 2017. Harnessing Effector-Triggered Immunity for Durable Disease Resistance. *Phytopathology,* pp. PHYTO-03-17-0086-RVW.

Zhang, S. and Klessig, D.F., 2001. MAPK cascades in plant defense signaling. *Trends in Plant Science,* 6(11), pp. 520-527.

Zhang, Y. and Ding, Y., 2020. The Simultaneous Coupling of Transcription and Splicing in Plants. *Molecular Plant,* 13(2), pp. 184-186.

Zhang, Y., Wang, Y., Liu, J., Ding, Y., Wang, S., Zhang, X., Liu, Y. and Yang, S., 2016. Temperature-dependent autoimmunity mediated by *chs1* requires its neighboring *TNL* gene *SOC3*. *New Phytologist,* 212(2).

Zhang, Y.-M., Shao, Z.-Q., Wang, Q., Hang, Y.-Y., Xue, J.-Y., Wang, B. and Chen, J.-Q., 2016. Uncovering the dynamic evolution of nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes in Brassicaceae. *Journal of Integrative Plant Biology,* 58(2), pp. 165-177.

Zhao, M., He, L., Gu, Y., Wang, Y., Chen, Q. and He, C., 2014. Genome- Wide Analyses of a Plant- Specific LIM- Domain Gene Family Implicate Its Evolutionary Role in Plant Diversification. *Genome Biology and Evolution,* 6(4), pp. 1000-1012.

Zhao, T., Rui, L., Li, J., Nishimura, M.T., Vogel, J.P., Liu, N., Liu, S., Zhao, Y., Dangl, J.L. and Tang, D., 2015. A Truncated NLR Protein, TIR-NBS2, Is Required for Activated Defense Responses in the exo70B1 Mutant. *PLOS Genetics,* 11(1), p. e1004945.

Zhou, J.-M., Tang, D. and Wang, G., 2017. Receptor kinases in plant pathogen interactions: more than pattern recognition. *The Plant Cell.*

Zhu, S., Li, Y., Vossen, J.H., Visser, R.G.F. and Jacobsen, E., 2012. Functional stacking of three resistance genes against Phytophthora infestans in potato. *Transgenic Research,* 21(1), pp. 89-99.

Zipfel, C. and Oldroyd, G.E.D., 2017. Plant signalling in symbiosis and immunity. *Nature,* 543(7645), pp. 328-336.

Zou, B., Sun, Q., Zhang, W., Ding, Y., Yang, D.-L., Shi, Z. and Hua, J., 2017. The Arabidopsis Chromatin-Remodeling Factor CHR5 Regulates Plant Immune Responses and Nucleosome Occupancy. *Plant and Cell Physiology,* 58(12), pp. 2202-2216.

Üstün, S., Sheikh, A., Gimenez-Ibanez, S., Jones, A., Ntoukakis, V. and Börnke, F., 2016. The Proteasome Acts as a Hub for Plant Immunity and Is Targeted by *Pseudomonas* Type III Effectors. *Plant Physiology,* 172(3), pp. 1941-1958.